A Publication of the American Psychological Association

The Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

Author(s)
"Complexity-Simplicity as a Personality Dimension"
Frank Barron . . .

A "Sleeper" Effect in Opinion Change . . .
Walter Weiss

Social Perceptions of Authoritarians and Nonauthoritarians . . .
Morris Scodel and Paul Mussen

The Attitudes of the Mothers of Male Schizophrenics toward Their Children . . .
Joseph C. Mark

Personality Differences between Persons Classified as Psychoneurotic and as Nonpsychoneurotic . . .
Leonard Eron

Effects of Initial Serial Position and Attitude upon Attitude Change under Conditions of Low Motivation . . .
Leonard W. Doob

Relation of Background Experience to Social Acceptance . . .
Mortimer R. Feinberg

Continued on inside front cover

Vol. 48 APRIL 1953

CONTENTS—*Continued*

ARTICLES

Is There a Mechanism of Perceptual Defense? 233
Leo Postman, Walter C. Bronson, and George L. Gropper

A Cross-Cultural Investigation of Closure 267
Donald M. Michael

Sharing Leadership in Small, Decision-Making Groups 291
Leonard Berkowitz

Concept of "Subsidy" in Sentence Psychiatric and Personality Tests: An Experimental Approach 339
William Seeman

Manifest Anxiety and "Perceptual Defense" 348
M. E. Bitterman and Calvin W. Kniffin

Neurosis and Tension: An Exploratory Study of the Relationship of Physiological and Rorschach Measures 353
Julius Wishner

On the Theory of Psychoanalytic Displacement 362
Robert R. Bush and John W. M. Whiting

Small Group Discussions with Participatory and Supervisory Leadership 273
A. Paul Hare

The Influence of Individual Members on the Characteristics of Small Groups 276
William Haythorn

A Personality Scale of Manifest Anxiety 285
Janet A. Taylor

Traumatic Avoidance Learning: The Outcomes of Several Extinction Procedures with Dogs 291
Richard L. Solomon, Leon J. Kamin, and Lyman C. Wynne

CASE REPORT

Reflections before the Gallows: The Last Testament of Jerome Titus 303
Percy Black

COMMENTS AND NOTES

The Ethnic Prejudices of White and Negro College Students . . . 311
J. Stanley Gray and Anthony H. Thompson

The Basic Law of Rumor 319
A. Chorus

REVIEWS 315

NEW BOOKS RECEIVED 325

Volume 48 APRIL 1953 Number 2

The Journal of Abnormal and Social Psychology

Editor: J. McV. Hunt, University of Illinois

Associate Editor: Arthur Jenness, Williams College

Book Review Editor: Stanley G. Estes, Harvard University

Managing Editor: Lorraine Bouthilet

CONSULTING EDITORS

Gordon W. Allport
Harvard University

Hadley Cantril
Princeton University

Lee J. Cronbach
University of Illinois

John Dollard
Yale University

Leonard W. Doob
Yale University

Charles M. Harsh
U. S. Navy Electronics Laboratory
San Diego

William A. Hunt
Northwestern University

W. Ernst Kris
Yale University School of Medicine

Donald W. MacKinnon
University of California

Quinn McNemar
Stanford University

O. Hobart Mowrer
University of Illinois

Theodore M. Newcomb
University of Michigan

Charles E. Osgood
University of Illinois

Saul Rosenzweig
Washington University

Harold Seashore
Psychological Corporation

David Shakow
University of Illinois School of Medicine

Ross Stagner
University of Illinois

Elizabeth Collins Wesman
White Plains, N.Y.

Robert W. White
Harvard University

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

PUBLICATION OFFICE: 372 BROADWAY, ALBANY, N.Y., AND

1333 SIXTEENTH STREET N.W., WASHINGTON 6, D.C.

Entered as second-class matter, June 18, 1923, at the Post Office at Albany, New York, under the Act of February 28, 1925. Acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, authorized August 26, 1947.

Copyright 1953 by the American Psychological Association, Inc.

The JOURNAL is issued quarterly in January, April, July, and October. The yearly volume averages 544 pages. The subscription price per volume is $7.00; foreign, $7.50. Single issues are $2.00.

Original manuscripts and editorial communications should be sent to:

DR. J. MCV. HUNT
Department of Psychology
University of Illinois
Urbana, Ill.

Books for review should be sent to:

DR. STANLEY G. ESTES
Psychological Clinic
64 Plympton Street
Cambridge 38, Mass.

Fifty free offprints are supplied to all contributors of articles. For book reviews, five copies of the JOURNAL are supplied gratis if requested.

Printed and mailed by:

THE BOYD PRINTING COMPANY, INC.
372 Broadway, Albany 7, N.Y.

Subscriptions and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC
1333 Sixteenth Street N.W.
Washington 6, D.C.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

COMPLEXITY-SIMPLICITY AS A PERSONALITY DIMENSION[1]

FRANK BARRON

Institute of Personality Assessment and Research, University of California

THIS is the fourth in a series of reports on the identification and measurement of a factor in perceptual preferences which seems to have considerable generality in human behavior. It may be described briefly as a bipolar factor which opposes a preference for perceiving and dealing with complexity to a preference for perceiving and dealing with simplicity, when both of these alternatives are phenomenally present and when a choice must be made between them. The kinds of phenomenal fields which have been the objects of the research to be reported here, and which seem to possess discriminable degrees of simplicity and complexity, are interpersonal relations and innerpersonal psychodynamics, politics and economics, religion, relations to authority, attitude towards sensual experience, social conformity and adherence to tradition, and originality and intellectual independence.

As we shall point out more fully later, a somewhat similar factor had previously been identified in the field of esthetic preferences by Eysenck, who attempted also to construct a measure of it (4, 5). It was, however, rediscovered independently and more or less incidentally by Welsh, who published the first of the four reports referred to above (12). A new measure of it was developed by Welsh and the present writer (3). The latter explored some of its correlates in preferences for painting and in self-conception (2). In the present paper we shall deal with further correlations of this dimension with variables in several areas of behavior. Before this further work is reported, however, a brief review of the findings which led up to it seems in order.

THE WELSH FIGURE PREFERENCE TEST

This series of investigations began with the construction by Welsh of a test consisting of several hundred line drawings in black ink on 3×5-in. white cards. He intended it as a nonverbal psychiatric diagnostic instrument, and sought originally to develop scales to measure such variables as Hysteria, Depression, Schizophrenia, Paranoia, and so on. The subject (*S*) was asked to indicate for each figure whether he liked it or did not like it, and thus to sort the drawings into two groups according to his preferences.

The effort to develop a diagnostic instrument is still in progress, but did not immediately meet with success. In seeking to understand the nature of the preferences being expressed, however, Welsh carried out a factor analysis of various scales defined on the basis of his own judgment as to the stimulus-character of the drawings: e.g., bilaterally symmetrical figures, three-dimensional figures, figures with many projections, figures with few projections, ruled-line drawings, free-hand drawings, angular figures, curved figures, and so on.

From this analysis, two factors emerged: an acceptance-rejection factor (expressing the general tendency of the subject either to like or to dislike the figures), and a second, bipolar factor, orthogonal to the first, whose poles, as determined by inspection of the figures, seemed to be simplicity (combined with a rather obvious, bilateral symmetry) and complexity (usually associated with a much less obvious kind of balance, which in previous reports has been referred to as "asymmetry").

This latter factor bears a close resemblance to that earlier identified by Eysenck, and named by him the K factor. Eysenck has demonstrated for a number of stimulus classes (colors, odors, paintings, polygons, poetry) the existence not only of a general factor of esthetic appreciation, but of a secondary, bipolar factor as well. The second factor presents the same polar opposition noted by Welsh, one pole being represented by preference for the simple polygon, the strong, obvious odor, the poem with the obvious rhyming scheme and the definite, unvarying, simple rhythm, and the simple, highly unified picture; at the other pole is

[1] This study was carried on at the Institute of Personality Assessment and Research, University of California, and was supported financially by the grant of the Rockefeller Foundation to that Institute. The writer owes a special debt of gratitude to the Director of the Institute, Dr. Donald W. MacKinnon, for his encouragement of the research and for his critical reading of the manuscript.

preference for the more complex polygon, the more subtle odors, the poem with a less obvious rhythm and a more variable and loose rhyming scheme, and the complex, more diversified picture.

THE BARRON-WELSH ART SCALE

A measure of this secondary, bipolar factor in the Figure Preference test was constructed in much the same incidental manner as that in which the factor had been discovered. It had happened that several artists were included in the control sample which Welsh had used for comparison with psychiatric patients, and these artists all clustered together at the complex-asymmetrical pole of the factor. The present writer, in search for measures of artistic discrimination for inclusion in a battery of assessment procedures, recalled this finding, and was led to wonder whether the factor was not significantly related to ability to discriminate the good from the poor in artistic productions. In any case, the Figure Preference test clearly consisted of stimulus material which might yield such a measure. The most straightforward way of checking on this seemed to be to give the entire 400-item test to a sample of artists and nonartists, and then to construct, by means of the item-analysis technique, a scale which would embody the differences between artists and nonartists in their preferences for the figures; and finally, of course, to cross validate the scale on new samples of artists and nonartists. This was accordingly done, and a highly reliable and valid scale resulted (3).

Inspection of the items in the scale quickly revealed that the secondary, bipolar factor which had emerged from the factor analysis was reproduced in the later empirically derived measure. The artists *liked* figures which were highly complex, asymmetrical, free-hand rather than ruled, and rather restless and moving in their general effect. (Several artists, in reacting to them, had described them as "organic.") The figures which were *liked* by people in general, however, were relatively simple, often bilaterally symmetrical, and regularly predictable, following some cardinal principle which could be educed at a glance. These figures were described by artists as "static," "dull," "uninteresting."

This convergence of the results of factor-analytic and external-criterion methodologies seems especially worth noting. Factor analysis not only revealed the psychological unity in perceptual preferences with which we are here dealing, but in addition provided the clue to an extremely important external correlate which could be used for straightforward empirical scale derivation. The scale is now properly designated as an Art Scale, but can be equally properly construed as a measure of the factor found by Eysenck and by Welsh. It has both psychological unity and external predictive power. Most of the remainder of this paper will be devoted to what Cattell refers to as "peripheral validation," which consists essentially of an extension of the investigation to correlates of the factor in areas of behavior which are, prima facie, remote from that in which the factor was discovered. First, however, we shall review briefly some previously reported findings concerning the relationship of these figure preferences to preferences in paintings and to self-descriptions.

ART PREFERENCES AND SELF-DESCRIPTION

The two test procedures which have already been analyzed and reported on (2) in relation to the Art Scale are the Gough Adjective Check-List (8) and a Painting Preference test assembled by the present writer. The former is a list of 279 common personally descriptive adjectives, from among which *S* is to select (by checking) those which he thinks describe himself; the latter consists of 105 postcard-size reproductions in color of paintings by a large number of European artists, which *S* is asked to sort into four groups according to the degree of his liking for the paintings.

The polarity noted in these three domains of figure preferences, art preferences, and adjective self-descriptions may best be summed up in this fashion:

I

In Figure Preferences

Preferring what is simple, regularly predictable, following some cardinal principle which can be educed at a glance.

In Art Preferences

Preferring themes involving religion, authority, aristocracy, and tradition.

In Adjective Self-Checks

Contented, gentle, conservative, patient, peaceable, serious, individualistic, stable, worrying, timid, thrifty, dreamy, deliberate, moderate, modest, responsible, foresighted, conscientious.

II

In Figure Preferences

Preferring what is complex, irregular, whimsical.

In Art Preferences

Preferring what is radically experimental, sensational, sensual, esoteric, primitive, and naive.

In Adjective Self-Checks

Gloomy, pessimistic, bitter, dissatisfied, emotional, pleasure-seeking, unstable, cool, irritable, aloof, sarcastic, spendthrift, distractible, demanding, indifferent, anxious, opinionated, temperamental, quick.

FURTHER RESULTS

The relationships cited in the preceding section were found in a sample of 40 male graduate students, from about a dozen departments of the University of California, who had taken part in weekend living-in assessments at the Institute of Personality Assessment and Research. These *S*s had been carefully studied by means of a large number of objective tests and experimental procedures, and had also been interviewed concerning their personal history, their values and philosophical views, their professional aspirations, and so on. At the conclusion of the series of (four) assessments, they were all rated by the staff [2] on 40 variables which previous personality researches had indicated as being of general importance.

We are thus in a position to see how the simplicity-complexity dimension is manifested in many different areas of behavior, by correlating the Figure Preference measure with other variables. For the most part, we shall restrict ourselves to relationships obtaining in this one sample of 40 *S*s. However, the scale was also given, together with some questionnaires, to a sample of over 100 male undergraduates in two Pennsylvania colleges, and the correlates of Complexity in that sample will be described. Data are also available from the assessment study of 80 other subjects (PhD candidates and medical school seniors) who took a revised form of the test.

For ease of reporting, we shall adopt the technique of drawing a composite picture of two ideal, modal persons, the simple and the complex, on the basis of the correlational results. The statistical support for the portrait will be cited as we go along. Several cautions must be observed in interpreting such material, however. For one thing, since the picture is based entirely on group relationships, it will fail in some respects to do justice to unique patterning of the variables in individual cases. Like the average man, the composite simple or complex person would be hard to find.

In addition, one must be particularly cautious in evaluating the "goodness" and "badness" of the correlates of this dimension. As a previous report (2) pointed out, it is important to bear in mind that, in terms of the total constellation of factors making for personal effectiveness and professional promise, simple persons and complex persons were equally represented among *S*s who were rated as possessing that combination of attributes in high degree. This equal representation held also among the group of *S*s with low ratings. One must conclude that both simplicity and complexity have their effective and ineffective aspects; they simply result in different sorts of merits and liabilities.

As we have indicated, in order to facilitate the reporting of the data, we shall adopt two conventions, one with regard to the scale and the other with regard to the designation of the *S*s. The Barron-Welsh Art Scale will hereafter be referred to as a measure of the variable *Complexity*, since it is that feature of the scale which is of interest here, and since the scale is so scored that preferences like those of artists (hence, preference for the complex) earn *S* a high score, while preferences like those of people in general (i.e., preference for the simple) earn a low score. The designations *Complex person* and *Simple person* will be employed to indicate a modal high scorer and a modal low scorer, respectively, on this particular test.

THE CORRELATIONAL COMPOSITE

This description begins with personal tempo, which is usually a rather easily observed, surface attribute, complicated though its ramifications may be in the personal character. It is "surface" in the sense that it is what we are first presented with when we meet another person; we take in almost automatically such attributes as flow of speech, speed of response, rate and intensity of expressive movement, and expansiveness or con-

[2] Staff members who participated in these assessments were D. W. MacKinnon, R. N. Sanford, Erik H. Erikson, R. S. Crutchfield, R. E. Harris, H. G. Gough, P. Dempsey, R. Taft, and the present writer. It need hardly be said that this entire research project is greatly indebted to the staff who contributed the ratings and carried out the assessment.

striction in interaction with the environment.

It will be recalled that on the Adjective Check-List our Complex *Ss* had described themselves as "quick" and "temperamental." The Simple *Ss*, on the other hand, had characterized themselves as "deliberate" and "dreamy." These self-descriptions would seem to be borne out by the staff's ratings of *Ss* on such variables as Personal Tempo, Verbal Fluency, and Constriction. The correlations with Complexity are, respectively, .50, .29, and —.42.[8] Thus, the Complex person is more intensely expressive, expansive, and fluent in speech than the Simple person. The Simple person, on the other hand, is seen as being more natural and likeable, and also as more straightforward and lacking in duplicity. (Complexity correlates —.44 with Naturalness, —.27 with Likeability, and .56 with Deceitfulness, as rated by the staff.) This picture of easy and uncomplicated simplicity is further supported by staff ratings of such factors as Good Judgment, Adjustment, and Abundance Values, all of which go with preference for the simple figures. (The *r*'s with complexity are —.39, —.31, and —.34, respectively.)

"Adjustment" had here been defined as "getting along in the world as it is, adequate degree of social conformity, capacity to adapt to a wide range of conditions, ability to fit in." As we shall see later, this kind of adjustment is not an unmixed blessing; the "unadjusted" complex person, who does not fit in very well in the world as it is, sometimes perceives that world more accurately than does his better-adjusted fellow.

The negative correlation with Abundance Values, combined with the positive relationship of Complexity to Deceitfulness, merits some comment. Abundance Values was defined as "Sense of security and optimism regarding the future, absence of fears of deprivation, of being exploited, and of being cheated." Deceitfulness was identified with "duplicity, lack of frankness, guile, subterfuge." Again, one recalls the adjective self-descriptions of the Complex people: gloomy, pessimistic, bitter, dissatisfied, demanding, pleasure-seeking, spendthrift. There is certainly some suggestion here of early oral deprivation, of pessimism concerning the source of supply, which is seen as untrustworthy and which must be coerced, or perhaps tricked, into yielding. It is as though the person had reason to believe that he would not "get what was coming" to him unless he made sure that he did, by whatever device might be available. It is this lack of infantile *trust* (as Erikson names it) that leads to adult duplicity and craftiness. One aspect of complexity then (and perhaps a penalty sometimes attaching to it) is, to render it in the common phrase, a sort of "two-facedness," an inability to be wholly oneself at all times. The more simple, natural, and likeable person finds it easier to be always himself.

This suggested relationship of Complexity with the derivatives of orality (i.e., character traits determined in part by a relatively long and intense oral stage of development in the child) receives some slight support from several other correlations. If it is true, as clinical evidence generally indicates, that oral fixation leads to feminine character traits in men, one would expect to find Complexity related to masculinity-femininity measures. This proves to be the case, the *r* with the Minnesota Multiphasic Personality Inventory *Mf* being .18, that with the Strong Interest inventory *Mf* (which is so scored that the feminine man earns a low score) —.39. Rated Effeminacy, defined as "Effeminate style and manner of behavior; softness," is also related to Complexity, the *r* being .29. Thus the Complex man has an attitude of acceptance towards his femininity, showing itself in soft, gentle, and effeminate behavior. In the light of this, it is not surprising to find both Sentience ("seeking and enjoying sensuous impressions, sensitive, esthetic") and Sensuality ("acceptance of and capacity for sensual gratification") being related positively to Complexity (*r*'s of .25 and .26 respectively). This completes a picture of low but consistent correlations of Complexity with some of the derivatives of orality.

The preference for Complexity is clearly associated with originality, artistic expression, and excellence of esthetic judgment. Originality was one of the three criterion variables around which the assessment research program was organized, and every subject was rated by the faculty members of his department on the degree of Originality he had

[8] All correlations reported in this paper are based on *N*'s of 40, with the standard error therefore being .16.

displayed in his work. The Complex person is seen as more original, both by the assessment staff and by the faculty of his department. The correlation with the criterion ratings on Originality is .30.

Complexity is also related to Basic Good Taste as measured by a test which presents various alternative arrangements of formal design elements and asks *S* to choose the most esthetically pleasing combination. This test, constructed by Sanford E. Gerard, is scored a priori in terms of known principles of composition. The correlation with Complexity is .44.

The *Ss* were also given a mosaic construction test devised by Turney, and known as the Turney Designs. In this test, the task is to construct a mosaic design in rectangular form from several hundred one-inch square solid-colored pieces of pasteboard (20 different colors being represented). The designs were then rated by members of the Art Department of the University, in terms of the artistic merit of the productions. The ratings thus obtained correlate .40 with Complexity.

To be purely speculative for the moment, one might wonder whether there is not some relationship between the more enduring and intense oral stage of development in our Complex *Ss* suggested by some of the data, and their evidently greater originality and sensitivity to the esthetic character of objects. Fowler (7) has shown that oral character traits are significantly associated with textural responses to shading on the Rorschach, and psychoanalytic writers, particularly Rank, have emphasized the relationship between femininity and artistic productiveness in men. Sanford (personal communication) has noted a tendency towards somewhat slower social development in the earlier years in original people, which would fit in nicely with the notion that a person must, as it were, have more commerce with himself and his feeling states and less with the environment during childhood if he is later to have sufficient communication with his own depths to produce the original thought. In this view, originality evidenced in maturity is to some extent dependent upon the degree to which the person in early childhood was faced with a complicated relationship to the maternal source of supply, combined with his capacity to persist at and eventually to achieve some mastery of this earliest problem situation. The argument would be that this primitive experience of phenomenal complexity sets a pattern of response which results in slower maturation, more tentativeness about the final form of organization, a resistance to early crystallization of the personality, and finally, greater complexity in one's view both of the outer and of the inner worlds.

Perhaps such speculation is unwarranted, however, and in any case it is clear that a great many other factors are involved in determining originality. What can be said is that originality and artistic creativeness and discrimination are related to the preference for complexity, and that the latter bears some relationship to sentience and femininity in men.

The Complex person's greater flexibility in thought processes is shown by a correlation of —.35 with rated Rigidity, defined as "inflexibility of thought and manner; stubborn, pedantic, unbending, firm." A 22-item scale for Rigidity, developed by Gough and Sanford, yielded an r of —.18 with Complexity. The two scales (Rigidity and Complexity) were later found to be related significantly and negatively to one another in a sample of medical students (r of —.36).

That repressive overcontrol may sometimes be associated with the preference for simplicity has already been indicated by the correlation of —.42 of Complexity with Constriction, and by another correlation of .50 with Impulsiveness. It is shown also in the relation of the Complexity measure to psychiatric variables which are scaled on the Minnesota Multiphasic Personality Inventory. With Hysteria, for example, Complexity correlates —.30, while with Schizophrenia it correlates .37, and with Psychopathic Deviate .36. Thus Complexity goes along both with lack of control of impulse (the *Pd* scale) and with the failure of repression which characterizes the schizophrenic process. This is by no means to suggest that any of these graduate students showed schizophrenic tendencies of a pathological degree, but it is reasonable to suppose the *Sc* scale of the MMPI has built into it the correlates of the sort of free-floating symbolic activity and frank confrontation and expression of the unconscious which is often so startlingly present in schizophrenic patients. The *Hy*

scale, on the other hand, picks up the tendency of *S* to repress aggressive and erotic impulses, or to render them innocuous by rationalization, reinterpretation, or gratification in a substitutive manner which will not cause conflict. At the risk of being oversimple, we might say that preference for the complex in the psychic life makes for a wider consciousness of impulse, while simplicity, when it is preferred, is maintained by a narrowing of that consciousness.

That the perceptual decision in favor of admitting complexity may make also for greater subjectively experienced anxiety is indicated by the correlation (.34) of Complexity with Overt Anxiety as measured by the Welsh Anxiety Index (11) on the MMPI. To tolerate complexity one must very often be able to tolerate anxiety as well, this finding would seem to say.

The negative correlation with *Hy*, and the positive correlation, of about the same magnitude, with *Sc* would seem to fit well with a finding of Eysenck's that his own measure of the Complexity-Simplicity dimension, the K test, correlates in this same direction with his Hysteria-Dysthymia factor.

Complexity is related negatively, although not significantly so, to Political-Economic Conservatism (—.22) and to Ethnocentrism (—.27), as measured by Form 60 of the Levinson-Sanford scale (10). These *r*'s are bound to be minimum estimates of the true relationship, since there was considerable restriction of range on both *PeC* and Ethnocentrism, the means being significantly lower than that of the general population. Some corroboration of this finding is furnished by attitudes expressed by *Ss* in the Pennsylvania sample, as we shall show later.

In addition to being less rather than more conservative, the person who prefers complexity is socially nonconformist. Staff ratings of Conformity correlate —.47 with Complexity, while the self-ratings of *Ss* on Conformity correlate —.53. Related to this is a correlation of —.29 between Complexity and Submissiveness, which was here defined as "deference, willingness to be led, compliance, over-ready acceptance of authority." In addition, Complexity correlates positively (.36) with the *F* scale of the MMPI. The *F* scale consists of items which are psychologically heterogeneous, the defining property of the scale being that all items in it have a low probability (about .1) of being answered in the scored direction. Thus, the higher the *F* score, the more likely it is that the subject holds a set of socially dissident and deviant opinions.

Complexity and Independence of Judgment

There is one further bit of information, of a rather intriguing nature, which bears on this question of conformity. Asch (1) has devised an experimental social situation in which *S* is put under pressure to conform to a group opinion which is false. In this experiment, there are from eight to sixteen ostensible *Ss*, only one of whom, however, is naive; the rest are in the hire of the experimenter. The task is to judge which of three lines of variable length meets a standard line; or, put otherwise, to match the length of a given line with one of three lines which are themselves not equal to one another. The *Ss*, one by one, announce their judgments publicly. The naive *S* is so placed as to be one of the last to announce his judgment. On the critical trials, the hired majority gives a prearranged false answer. The experimental variable is called Yielding, which is defined as agreeing with group opinion when it is in error. Yielding scores, in the prototypical experiment, range from 0 to 12, zero yielding being known as Independence.

In this experiment, approximately 25 per cent of all *Ss* are Independent, while the 25 per cent at the other extreme of the distribution Yield from 8 to 12 times. These two quartiles were selected for intensive personality study, in which the present writer participated as a test consultant and analyst. Among the tests employed in this part of the research was the measure of Complexity whose correlates we have been describing.

Since the experiment and the personality study of *Ss* will be fully reported elsewhere, we shall note here only that the mean difference in Complexity scores between the Independents and Yielders was in the predicted direction, i.e., Independents were more Complex. In a sample of 40 Haverford College students, the difference was significant at the .001 level, while in a sample of 60 Temple University students it was significant at the .05 level. For both groups combined, the difference was significant at the .01 level.

We have referred earlier to the negative relation of Complexity to Adjustment, part of the definition of which involved "an adequate degree of social conformity," and "ability to fit in." It is indeed a delicate question as to what, in social conformity, is

an adequate degree thereof; different people seem to make different requirements of others in this matter. It is almost implicit in the design of the Asch experiment that to abandon the evidence of one's own senses in favor of group opinion is carrying a good thing too far; while the popularity of many a modern nation within its own boundaries tends to become so categorical that even judicious dissent from this or that aspect of national policy may get labeled "disloyal," "deviationist" or, the shift in adaptation level in invective being what it is, something even worse.

In any event, it seems that these complex and independent *Ss* did not "fit in"; they faced up to the anxiety of being a minority of one, and continued to call the turn as they saw it. They accepted the fact that they held a different opinion from the rest of the group, and they were able to persist in their belief in their own opinion. Some, indeed, were not surprised at all at this, and seemed to take pleasure in being among the Opposition; others were made extremely anxious, to the point of panic; while there were other independent and complex *Ss* who, while somewhat perturbed, went their own way with unassertive but firm confidence in their own judgment. In this latter group, independence seemed more genuinely critical and differentiated; and it is this sort of independence which, when it is allied with complexity, should prove constructively unadjusted and actually productive of original work rather than simply a deviant attitude.

RESPONSES TO AN ATTITUDE QUESTIONNAIRE

In the same study, Asch and the present writer collaborated in the construction of a criterion-specific questionnaire, which consisted of 86 items which were especially selected, or written anew, to test particular hypotheses concerning personality differences between Independents and Yielders. Again, the results of this effort will be described in another publication; however, the results of an item analysis of the Independence questionnaire against the criterion of high or low score on Complexity may be given here. In this analysis, the Haverford and Temple groups were combined, and the questionnaire responses of the highest 27 per cent of scorers on the Complexity scale were compared with the lowest 27 per cent. Items which showed differences significant at the .05 level are given below.

Answered True by High Scorers on Complexity

1. The unfinished and the imperfect often have greater appeal for me than the completed and the polished.
2. I could cut my moorings . . . quit my home, my parents, and my friends . . . without suffering great regrets.
3. Politically I am probably something of a radical.
4. I think I take primarily an esthetic view of experience.
5. I would enjoy the experience of living and working in a foreign country.
6. Many of my friends would probably be considered unconventional by other people.
7. Some of my friends think that my ideas are impractical, if not a bit wild.
8. I enjoy discarding the old and accepting the new.
9. When someone talks against certain groups or nationalities, I always speak up against such talk, even though it makes me unpopular.

Answered True by Low Scorers on Complexity

1. I don't like modern art.
2. Disobedience to the government is never justified.
3. Perfect balance is the essence of all good composition.
4. Straightforward reasoning appeals to me more than metaphors and the search for analogies.
5. It is a pretty callous person who does not feel love and gratitude towards his parents.
6. Things seem simpler as you learn more about them.
7. I much prefer symmetry to asymmetry.
8. Kindness and generosity are the most important qualities for a wife to have.
9. When a person has a problem or worry, it is best for him not to think about it, but to keep busy with more cheerful things.
10. It is the duty of a citizen to support his country, right or wrong.
11. Barring emergencies, I have a pretty good idea what I'll be doing for the next ten years.
12. I prefer team games to games in which one individual competes against another.
13. An invention which takes jobs away from people should be suppressed until new work can be found for them.

Since these *Ss* were all undergraduate students in Pennsylvania colleges, and their average age was 19, it is of some interest to compare their characteristics with the correlates of Complexity noted in the sample of California PhD candidates. By grouping items, we may summarize the personality differences between the Complex and Simple persons in the Pennsylvania sample somewhat as follows:

The Complex person:

1. is artistic (4);
2. has unconventional friends, occasionally is visited by an impractical, not to say wild, idea, and would rather be creative and neurotic than normal and ordinary (6,7,10);
3. is politically somewhat radical, and can be militantly opposed to racial prejudice (3,9);
4. is aware of present imperfections, would welcome and has faith in future developments (1,2,5,8).

The Simple person:

1. doesn't like modern art (1);
2. particularly values kindness and generosity in a wife (as opposed to implied alternative values), and feels that the proper filial sentiments towards one's parents are love and gratitude (8,5);
3. feels that a citizen should support his country, right or wrong, and that disobedience to the government is never justified. Somewhat allied to this, he prefers a team effort to individual competition (2,10,12);

4. prefers symmetry to asymmetry, considers perfect balance the essence of good composition, and prefers straightforward reasoning to metaphors and the search for analogies (7,3,4);

5. has clear plans for the future, and considers that things seem simpler as you learn more about them (11,6);

6. believes that a person with a problem or worry should not think about it, and that inventions which take jobs away from people should be suppressed until new work can be found for them (9,13).

The general pattern is certainly similar to that shown by the correlations we have been reporting. In Pennsylvania as in California, preference for simplicity is associated with social conformity, respect for custom and ceremony, friendliness towards tradition, somewhat categorical moral judgment, an undeviating patriotism, and suppression of such troublesome new forces as inventions which would temporarily cause unemployment. This last item is almost prototypical of the simple person's orientation towards repression as a psychic mechanism. In the California sample, it was shown in the negative correlation of Complexity with Hysteria and the positive correlations with Psychopathic Deviate and Schizophrenia. Its derivatives appear in many other characteristic attitudes as well, such as acceptance or rejection of sensual experience, of conventional religion, of paintings of unclad ladies, and so on. For "invention," write "impulse," and it is not hard to see an analogy to the common clinical formulation of the function of repression in the hysterical character.

The correlates of preference for complexity in this undergraduate sample are, again, much like the correlates of the corresponding preference among the PhD candidates. Complexity goes along with artistic interests, unconventionality, political radicalism, strong cathection of creativity as a value (even at the expense of "normality," as the item puts it), and a liking for change.

It seems evident that, at its best, preference for simplicity is associated with personal stability and balance, while at its worst it makes for categorical rejection of all that threatens disorder and disequilibrium. In its pathological aspect it produces stereotyped thinking, rigid and compulsive morality, and hatred of instinctual aggressive and erotic forces which might upset the precariously maintained balance.

There is a passage in Hugo's *Les Miserables* which is remarkably coincident with these observations. It occurs at that point in the narrative when Javert, the single-minded and merciless representative of the law, has turned his own world upside-down by allowing Jean Valjean, the outlaw whom he had so relentlessly pursued, and whom he finally had in his grasp, to escape. He says to himself, in this surprising moment, "There is something more then than duty." At this, "he was startled; his balances were disturbed; one of the scales fell into the abyss, the other flew into the sky. . . ."

> . . . To be obliged to acknowledge this: infallibility is not infallible, there may be an error in the dogma, all is not said when a code has spoken, society is not perfect, authority is complicate with vacillation, a cracking is possible in the immutable, judges are men, the law may be deceived, the tribunals may be mistaken . . . to see a flaw in the immense blue crystal of the firmament!
>
> . . . Certainly it was strange, that the fireman of order, the engineer of authority, mounted upon the blind iron-horse of the rigid path, could be thrown off by a ray of light! that the incommutable, the direct, the correct, the geometrical, the passive, the perfect, could bend!
>
> . . . Until now all that he had above him had been in his sight *a smooth, simple, limpid surface; nothing there unknown, nothing obscure; nothing which was not definite, coordinated, concatenated, precise, exact, circumscribed, limited, shut in, all foreseen;* authority was a plane; no fall in it, no dizziness before it. Javert had never seen the unknown except below. *The irregular, the unexpected, the disorderly opening of chaos,* the possible slipping into an abyss; that belonged to inferior regions, to the rebellious, the wicked, the miserable.[4]

This passage brings together many observations made intuitively by Hugo and arrived at in a more pedestrian manner in this research. A precise simplicity is seen to be related to authority, dogma, tradition, morality, constriction, and repression; the opposite of all these things is typified by the flaw in the crystal, by the irregular, by disorderly chaos; such qualities as are to be found in the inferior regions, where reside the rebellious, the wicked, and the miserable. The emphasis here is pathological, and the dichotomy absolute, but if we extend the range into normal behavior and admit the many shortcomings of the typology, there is considerable agreement between Hugo's intuition and this set of correlations.

[4] Italics mine.

RESERVATIONS AND DISCREPANT FINDINGS

So far as the correlational results in the California sample are concerned, it is evident that they must be accepted with some caution, since there is considerable possibility that chance error alone would account for some of the significant correlations. In addition, some variables would be correlated with one another in this sample but not in the general population, and a correlation of Complexity with just one of a set of atypically related variables would contribute to misleading correlations of the others of that set with the Complexity variable.

However, the findings from the Temple and Haverford nonassessment samples would tend to increase our confidence that many of the correlations found in the California assessment group would stand up when subjected to further investigation. There is also another source of evidence, although it is open to some question. The correlates of a revised and significantly different form of the Complexity measure in two other assessment samples (40 more PhD candidates, and 40 medical school seniors) are known. In general, they are very much the same as those reported in this paper; in some cases, the *r*'s are higher (—.39 with *PeC*, e.g., and .45 and .39 with Originality). However, there are some discrepant findings which should be noted, even though the Complexity measures are not strictly comparable.

In the first place, the negative relations of Complexity to Naturalness, Likeability, and Adjustment become zero-order in both the medical and graduate student samples. In addition, the positive relationship with Deceitfulness disappears, being slightly negative in the second sample of graduate students and exactly zero in the medical sample.

Several relationships which were zero in the first California sample become significantly positive in both of the succeeding samples. Complexity shows correlations with Intellect of .42 and .41; with Breadth of Interest of .33 and .39; with Sense of Humor of .39 and .33; with Cathexis of Intellectual Activity of .42 and .29. (The correlations in the graduate school sample are given first, those in the medical sample second.)

Work is going forward on further investigation of these results. Within a year or so a sample of at least 200 *Ss* will have been assessed, which should yield fairly stable correlations. An effort is being made to develop other measures of the complexity-simplicity dimension, with some attention to capacity as well as preference for dealing with and resolving complex phenomenal fields. The scale is also being used in nonassessment settings, such as psychiatric clinics, and as stimulus material in memory experiments.

SUMMARY

This paper has described the correlates of a bipolar factor in perceptual preferences which opposes a preference for perceiving and dealing with complexity to a preference for perceiving and dealing with simplicity, when both of these alternatives are phenomenally present and a choice is made between them. The correlates ramify through many areas of human behavior and attitudes, including innerpersonal psychodynamics, interpersonal relations, and more broadly social spheres of behavior, such as politics, religion, group interaction, and the like.

It was emphasized that these types have both their effective and ineffective aspects, so far as human functioning is concerned. At times there is considerable merit in the simple view, while on other occasions some ease may profitably be sacrificed for greater phenomenal richness.

A measure of this factor was described, and its correlates were reported. Calling the factor variable *Complexity*, we note this pattern of relationships to it:

1. It is related positively to personal tempo, verbal fluency, impulsiveness, and expansiveness.

2. It is related negatively in one sample to naturalness, likeability, lack of deceitfulness, adjustment, and abundance values, but in other samples a revised form of the measure shows no significant relationship to these variables, so that the finding must await further checking before being credited.

3. It is related positively to originality, good taste, and artistic expression, and its revised form in two other samples shows significant positive correlations with intellect, sense of humor, breadth of interest, and cathexis of

intellectual activity (none of which were significantly related to it in the first sample).

4. It is related positively to sensuality, sentience, esthetic interest, effeminacy, and femininity in men.

5. It is related negatively to rigidity and constriction.

6. It is related negatively to control of impulse by repression, and positively to expression of impulse and to breakdown of repression.

7. It is related negatively to political-economic conservatism, to subservience to authority, to ethnocentrism, and to social conformity.

8. It is related positively to independence of judgment.

REFERENCES

1. ASCH, S. E. Effect of group pressure upon the modifications and distortion of judgments. To appear.
2. BARRON, F. Personality style and perceptual choice. *J. Pers.*, in press.
3. BARRON, F., & WELSH, G. S. Artistic perception as a factor in personality style: Its measurement by a figure-preference test. *J. Psychol.*, 1952, 33, 199–203.
4. EYSENCK, H. J. The general factor in aesthetic judgments. *Brit. J. Psychol.*, 1941, 31, 94–102.
5. EYSENCK, H. J. "Type"-factors in aesthetic judgments. *Brit. J. Psychol.*, 1941, 31, 262–270.
6. EYSENCK, H. J. *Dimensions of personality.* London: Routledge, Kegan Paul, 1947.
7. FOWLER, C. Personality correlates of the differential use of shading on the Rorschach test. Unpublished bachelor's thesis, Bennington College.
8. GOUGH, H. G. Predicting success in graduate training. A progress report. Berkeley: Univer. of California, Institute of Personality Assessment and Research, 1950. Hectographed.
9. HUGO, V. *Les Miserables.* New York: The Modern Library.
10. LEVINSON, D. J., & SANFORD, R. N. A scale for the measurement of anti-Semitism. *J. Psychol.*, 1944, 17, 339–370.
11. WELSH, G. S. An anxiety index and an internalization ratio for the MMPI. *J. consult. Psychol.*, 1952, 16, 65–72.
12. WELSH, G. S A projective figure-preference test for diagnosis of psychopathology: 1. A preliminary investigation. Unpublished doctor's thesis, Univer. of Minnesota, 1949.

Received March 31, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

A "SLEEPER" EFFECT IN OPINION CHANGE [1, 2]

WALTER WEISS
Boston University

UNTIL recently the results of investigations in the area of attitude-opinion change have supported the assumption that the maximum modification of opinions is to be expected shortly after exposure to experimental treatment (1, 9, 12, 14). A finding of increased effects of a communication after a period of time has, however, been reported in an experiment carried out by the Information and Education Division of the War Department (5). The effect of a documentary film on certain opinions held by soldiers was found to diminish over a period of nine weeks, while an unexpected increment of effect from one to nine weeks was found for other opinions. One of the hypotheses offered by the investigators to account for this phenomenon was based on a distinction between the learning and acceptance of a communication. It stated that, although the content of a communication is learned well, the communication may be discounted as coming from a source having a propagandistic purpose. Over a period of time, the discounting factor may be forgotten more rapidly than the content on which the opinions are based. The result would be a delayed increment of the effect of the communication. Some of the implications of this hypothesis were confirmed in a recent study by Hovland and Weiss (6) on the influence of source credibility on the effectiveness of communications.

In both of the previous investigations the experimenters (*E*s) made the assumption that a discounting response inducing nonacceptance of the communication was aroused implicitly by the audience's reaction to the credibility and intentions of the source. The primary objective of the present research, therefore, was to test the prediction that a group influenced *explicitly* to discount the learned content of a communication would evidence a "sleeper" effect. The phenomenon would be considered demonstrated if following exposure to a communication the loss of its effectiveness with the passage of time is less for a group influenced by a discounting factor than for one not receiving this additional treatment. This interpretation of the "sleeper" effect encompasses the delayed increment of effects obtained in the previous studies as a particular case, depending on the coincidence of the time of opinion retest and removal of the discounting influence. Amplification of this point is made in the "Discussion" section.

In contrast to the procedures that had been employed, an attempt was made to control acquisition of the content. This was to insure that differential learning could not be construed as an explanation for a difference in the immediate or delayed effects on opinions. Also, in order not to prejudice the initial acceptance of the content, the discounting factor was introduced *after* the learning was completed. This represents a distinct change from the experimental situation that obtained in both earlier studies; for in those, the discounting motive may have been aroused and active during exposure to the communication, influencing acceptance of the content while the learning was occurring. The critical role played by nonacceptance of the communication would be more clearly shown if a sleeper effect were to be demonstrated under the present procedure.

The second major objective of the research pertains to the effect of opinions on the retention of content. A number of investigations have verified the broad hypothesis that a favorable attitude towards the viewpoint expressed in a communication leads to better retention of the content than does an unfavorable attitude (2, 4, 8, 11, 14, 15). Little or no attempt, however, was made to control

[1] This paper presents the principal parts of a dissertation submitted to the faculty of the graduate school of Yale University in candidacy for the degree of Doctor of Philosophy. The writer wishes to express his gratitude to Drs. C. I. Hovland, L. W. Doob, and F. D. Sheffield for their valuable advice and aid.

[2] This study was supported by funds from the Rockefeller Foundation for a coordinated research project on factors influencing changes in attitude and opinion being conducted at Yale University. The project is under the general direction of Dr. C. I. Hovland.

the acquisition of the material employed in these studies. It is possible, consequently, that the retention results could be explained in terms of differential learning. A further purpose of the current research was therefore to verify the validity of the general hypothesis when a measure of control is exerted over the learning of the communication content.

METHOD

Procedure

Three groups of experimental *Ss* were presented with a communication on "The Effects of Smoking," arranged in the form of a learning experiment. Their task was to learn to associate the word "true" or "false" with each of eight statements. The following instructions read to the students describe the procedure:

> A list of eight statements will be read to you. I will say after each statement whether it is true or false. You are to write nothing but just to listen. Then, the same list will be reread. But, now, you will not be told which are true and which false. This time, you will be given three seconds after each statement is read in which to remember whether the statement is true or false, and circle the word true or false on the prepared sheet. If you cannot remember, put a dash on that line.
>
> We will go through this same procedure four times.
>
> To show you how this will work, let us have a practice session with three statements (concerning presidents of the United States).
>
> As you probably noticed, the order of reading the statements changed each time. Therefore, do not try to memorize just the order of the words true and false. The only method that will work is to learn which statements are true and which false. Therefore, listen to the entire statement each time and try to remember whether it is true or false.

The actual statements used and labels applied were:

1. Heart disease occurs more frequently among tobacco smokers than among nonsmokers under 50 years of age. True
2. Fewer smokers are affected with hardening of the arteries than are nonsmokers. False
3. People who smoke have much less of a chance of reaching 60 than those who do not smoke at all. True
4. The heat of a lighted cigarette burns the delicate membrances of the nose, throat, and lungs, reducing the body's resistance to colds. True
5. The majority of smokers have normal blood pressure even after years of smoking. False
6. Ulcers, indigestion, and excess acid in the stomach are directly traceable to smoking. False
7. Studies have shown that smoking is an important factor in producing cancer of the lungs and throat. True
8. College students of today are not as tall as those of 20 years ago, because they smoke ten times as much. False

For one group (the "nondiscounted" group), this learning terminated the communication session. Another group (the "discounted" group) was given, immediately after, a brief countercommunication designed to arouse an attitude to discount or inhibit the acceptance of the prior learning. They were told that experimentation recently completed has resulted in evidence conclusively contradicting the findings of previous studies. Consequently, the truth or falsity of the statements they had just heard must be taken with a lot of skepticism and not as the last word on the matter.

The third experimental group (the "nondiscounted plus color-association" group) was presented with an additional learning task, which it was hoped might affect adversely their retention of the association. The same statements and procedure were employed again. But, instead of the words "true" or "false," they now had to learn to associate either the word blue, yellow, or orange with each of the eight statements.

The opinions of the *Ss* were assessed at least 10 days before the experimental sessions, immediately after, three weeks after, and six weeks after. Different subgroups from each of the three main experimental groups were retested at each of the three postexperimental time intervals. A retention trial was also included at each of these time intervals to determine whether the asserted truth or falsity of the statements comprising the communication was still remembered. The following instructions introduced the retention test:

> I will now read (just once) the statements which you listened to during the learning experiment. Some of them the speaker said were true and some false. You will be given three seconds after each statement is read in which to remember whether it is true or false, and circle the word true or false on the prepared sheet. If you cannot remember, put a dash on that line.

A group of control *Ss*, who did not receive any of the experimental treatments, was also employed. All took the before-questionnaire at the same time as the experimental *Ss*. Also, a different subgroup was retested at each of the three postexperimental intervals mentioned above.

RESULTS

Acquisition and Retention of Statement-Label Associations

Since acquisition was terminated on the fourth trial for all, it is necessary to determine if the mean amount learned differs significantly among the experimental groups.

An examination of Table 1 indicates that the *F* value for groups does not achieve the

TABLE 1

ANALYSIS OF VARIANCE OF MEAN NUMBER OF LABELS CORRECT ON LAST LEARNING TRIAL

SOURCE	*df*	SS	MEAN SQ.	*F**
Between groups	8	9.23	1.15	1.16
Within groups	210	208.11	0.99	
	218	217.34		

* *F* of 1.98 needed for .05 level of significance.

magnitude needed for significance. This indicates that the extent of acquisition of the eight statement-label combinations by the last learning trial does not differ beyond a chance level among the experimental subgroups. (The grand mean for acquisition was 7.51.) Consequently, differences among the experimental groups in opinion change cannot be considered a function of prior differences in the amount learned.

Although no data were procured with respect to variations in intelligence among the experimental groups, internal evidence suggests that the groups did not differ in learning ability. A detailed analysis of the acquisition data revealed no significant differences among the groups at the end of the first three trials.

Also, there are no significant differences among the treatment subgroups tested at the immediate and three-week periods. At the six-week interval, however, the subgroup given the color-association treatment remembers the statement-label combinations better than do the other two experimental subgroups.

On the whole, the color-associating treatment does not influence to a significant extent the retention of the statement-label associations. Accordingly, the two nondiscounted groups were combined into one enlarged nondiscounted category in the following analyses of the effect of the communication on opinions. The opinion results are in no way prejudiced by this procedure, for the two original nondiscounted groups were found not to differ in the immediate and retained effects of the communication.

TABLE 2

ANALYSIS OF VARIANCE OF MEAN NUMBER OF LABELS CORRECT ON RETENTION TEST

SOURCE	*df*	SS	MEAN SQ.	*F*	*F* NEEDED FOR .05 LEVEL
Treatment	2	3.39	1.70	1.10	3.04
Time	2	75.09	37.54	24.38	3.04
Interaction	4	21.93	5.48	3.56	2.41
Residual	210	324.26	1.54		
	218	424.67			

As noted in the section on "Procedure," a retention-test trial was included at each of the after-questionnaire periods. It is therefore possible to determine if the experimental treatments differentially affect retention of the learned associations. The appropriate analysis of the mean number retained is presented in Table 2.

The difference in the over-all effect of the treatments on retention is not significant. Differences between the groups in the loss of effectiveness of the communication may not therefore be considered a function of differences in retention of the learned associations.

Significant *F* values for time of retest and for the interaction between time and treatments are indicated. A further statistical examination of the means reveals a significant drop in retention for the groups tested at the three- and six-week intervals when compared with those tested immediately. (The retention means for the retest intervals are: immediate, 7.50; three weeks, 6.37; six weeks, 6.14.)

Immediate and Retained Over-All Effect of the Communication on Opinions

In order to determine the over-all effect of the learning on *Ss*' opinions on the topic of smoking, a mean change-score was devised for each individual.[3] This score was derived from the seven opinion items of the questionnaire on the topic of the communication. For each time interval of opinion retest, the mean group change-score for each experi-

[3] An individual's change-score for an opinion item was determined in the following manner. If *both* his before- and after-opinions were in agreement or disagreement with the position taken in the communication, he was assigned a zero score for that item. If his before-opinion was in agreement but his after-opinion was in disagreement, he received a score of −1. Finally, if his before-opinion was not in agreement and his after-opinion was in agreement, he was given a score of +1. The theoretical range is therefore 7 to −7. The use of an over-all score of this sort is necessary to determine the general effect of the communication on a person's opinions.

TABLE 3

MEAN OPINION-CHANGE SCORES FOR EXPERIMENTAL AND CONTROL GROUPS AT EACH PERIOD OF RETEST FOLLOWING THE COMMUNICATION SESSIONS

GROUP	CHANGE SCORE	*p* diff FROM CONTROL	*p* diff BETWEEN EXP. GROUPS
IMMEDIATE *			
Discounted (N=27)	.233	.0001	.09
Nondiscounted (N=47)	.289	.0001	
Control (N=25)	−.017		
AFTER THREE WEEKS **			
Discounted (N=21)	.136	.002	.45
Nondiscounted (N=52)	.102	.001	
Control (N=51)	−.022		
AFTER SIX WEEKS **			
Discounted (N=28)	.179	.005	.15
Nondiscounted (N=53)	.103	.09	
Control (N=38)	.041		

* *p* values are based on a one-tail test of significance.
** *p* values are based on a two-tail test of significance.

mental subgroup was assessed against that for the appropriate control subgroup. Also, the differences between the experimental groups were tested.

It is seen from Table 3 that learning significantly affects the opinions of both experimental groups in the direction of the communication. This effect not only appears immediately after the learning, but also persists over the six-week interval. An important implication of these data is that a further drop to the opinion baseline set by the control *Ss* could have been made by both experimental groups. Any differential loss of the effectiveness of the communication from the immediate to the delay testing periods is consequently not a spurious result of the lower level of immediate opinion-change shown by the discounted group.

The difference in mean opinion-change between the experimental subgroups at the immediate retest interval approaches significance. This indicates that the discounting treatment was immediately effective in reducing the extent of opinion-change evidenced by the discounted group. But since this group was still significantly above the control group, it is evident that the effect of the discounting treatment was not as great as that of the learning.

TABLE 4

DIFFERENCE BETWEEN MEAN OPINION-CHANGE SCORES FOR IMMEDIATE AND DELAY PERIODS OF RETEST FOR EACH EXPERIMENTAL SUBGROUP

GROUP	3 WEEKS MINUS IMMEDIATE	6 WEEKS MINUS IMMEDIATE	3+6 WEEKS MINUS IMMEDIATE
Discounted	−.097	−.054	−.073
Nondiscounted	−.187	−.184	−.185
Difference	.090	.130	.112
p *	.08	.03	.02

* *p* values are based on a one-tail test of significance.

The "Sleeper" Effect

As stated earlier, a sleeper effect would be considered demonstrated if the loss of the effectiveness of the communication were less for the discounted group than for the nondiscounted one. A statistical examination of this differential loss in changes of opinion is presented in Table 4.

The loss of the effect of the communication was not significantly less for the discounted group by the three-week interval. But at the longer delay period of six weeks, the differential loss of effectiveness was significantly less for that group. The use of a combined-delay category shows clearly the general tendency for the discounted group to lose less of the effectiveness of the communication within the time periods employed.[4] Thus, a sleeper effect is shown to have occurred.[5]

To provide evidence that the differential loss of effectiveness was not biased by the distributions of initial opinions, an analysis of the differences between experimental sub-

[4] A similar analysis was made based on the original three treatment subgroups—nondiscounted, discounted, and nondiscounted plus color association. The results are the same. The discounted group shows significantly less of a loss in the effectiveness of the communication at the six week and combined delay intervals than do either of the nondiscounted groups. The latter two do not differ significantly from each other.

[5] The statistical demonstration of the effect is based on group averages; as such, the phenomenon is demonstrated as a general group response and, consequently, need not necessarily be confirmed in the behavior of each individual. Personality factors, other related attitudes, differences in acceptance of the communication, etc. may be at the root of individual variations in behavior. To pinpoint more precisely the applicability of the communication-effect relationship, research on individual divergences from the average response is needed.

groups in before-opinion scores[6] was made. The mean scores for the subgroups retested at the same time period do not differ significantly from each other. The differences in mean before-opinion scores among the subgroups within the discounted and nondiscounted categories are also not significant.

An analysis of the differential loss of the effectiveness of the communication on the individual opinion items revealed that by the six-week period the discounted group showed less of a loss on each of the opinion questions. As noted above, it is at this retest interval that the analysis based on mean change-scores indicates a significant difference between the loss in communication effectiveness for the two groups.

counted and nondiscounted groups were combined. (A previous analysis revealed that they did not differ from each other.) An examination of this table shows that the statement-label association is retained better in those instances in which opinions were favorable rather than unfavorable before and after the communication. Similarly, a significant advantage in retention is seen for the category in which opinions were unfavorable before but favorable afterwards when compared with instances of unfavorable opinions before and after. These differences show up for the immediate retention test and increase for the delay testing periods. Although there are relatively few instances falling in the category of favorable opinion before and unfavorable

TABLE 5

PROPORTION OF TOTAL INSTANCES IN WHICH LABELS LEARNED CORRECTLY ON LAST LEARNING TRIAL WERE REMEMBERED ON RETENTION TEST, AS A FUNCTION OF BEFORE-AFTER CHANGE IN THE OPINION RELATED TO THAT LABEL

GROUP	IMMEDIATE			COMBINED DELAY			AVERAGE		
	N	PER CENT	*p**	*N*	PER CENT	*p**	*N*	PER CENT	*p**
Unfav. Before & After	112	89.39	.002	317	62.15	.0001	429	69.00	.0001
Fav. Before & After	165	97.58		314	88.85		479	91.86	
Fav. Before—Unfav. After	9	88.89	.12	49	77.55	.03	58	79.31	.01
Unfav. Before—Fav. After	126	98.41	.06	165	84.85	.23	291	90.72	.07
Unfav. Before & After	112	89.39	.002	317	62.15	.0001	429	69.00	.0001
Fav. Before—Unfav. After	9	88.89	.95	49	77.55	.04	58	79.31	.06

* *p* values are based on a two-tail test of significance.

Retention of Statement-Label Associations and Before-After Opinions.

Since six of the statement-label combinations bear an individual and direct relation to an opinion item, it is possible to determine whether the retention of the associations is dependent on before-after relations in opinions. By grouping separately *instances* of the four possible before-after opinion changes in relation to the retention of the appropriate associations, such an analysis can be made.

In Table 5 the data for the three- and six-week periods as well as those for the dis-

after, the evidence indicates a decided tendency for this group to yield generally poorer retention than either of the groups in which opinions were favorable after the communication. However, a comparison between the two categories in which after-opinions were unfavorable reveals an advantage in retention for the instances in which the before-opinions were favorable when compared with the instances in which the before-opinions were unfavorable.[7]

[6] An individual was given a score of one or zero for a questionnaire item, depending on whether his initial opinion agreed or disagreed with the position taken in the communication. The theoretical range of total scores is therefore from zero to seven. It is recognized that this procedure is somewhat arbitrary; but such a method does afford a means of obtaining a general opinion score based on persons rather than having to deal with the opinion items individually.

[7] It is possible that the instructions introducing the recall trial (see the section on "Procedure") received variant interpretations by the students. Instead of trying to recall the true or false labels applied by the experimenter to the statements, some students may have responded in terms of what *they* considered the factual status of the statements to be. Since there may be some ambiguity in the categorization of all the behavior as retention responses, it is appropriate to consider the qualifying phrase, "under the conditions of this experiment," when thinking about these data.

DISCUSSION

The current research had as its major objective the verification of two predictions: (*a*) the group given the discounting treatment would evidence a sleeper effect, and (*b*) the rate of forgetting of the associations would be differentially affected by the *Ss*' opinions. Both predictions were verified by the data. However, some of the results relative to the first prediction are in seeming disagreement with those of the study by Hovland and Weiss (6). Similarly, a full consideration of all the findings pertinent to the second prediction suggests an extension of the hypothesis from which it was derived.

The experimental investigation of Hovland and Weiss confirmed the existence of a sleeper effect by showing that communications ostensibly written by untrustworthy sources produced a greater net opinion-change after one month than immediately after, while the effect of those whose sources were regarded as trustworthy was less after one month than immediately after. This general result of a delayed enhancement of the effect of a communication is in line with the findings of the War Department study (5). In the present research, however, the effect of the communication on both experimental groups is greatest immediately after and decreases with the passage of time. A sleeper effect is evidenced only as a differential loss in the effectiveness of the communication from the immediate to the later retest periods.

The theory proposed by the previous investigators in explanation of the phenomenon is based on the presumed interactive influence on opinions of the learning and acceptance of communicated material. If the content is both learned and accepted, the maximum effect would probably be evidenced at once. If, however, a discounting motive interfering with acceptance of the learned content were aroused, the initial effect on opinions would be lowered. A delayed increment would occur if, with the passage of time, the interfering influence were decreased at a more rapid rate than the forgetting of the content on which the opinions are based. For the reduction in the extent of opinion-change due to forgetting of the content would be more than offset by the removal of the interference associated with the discounting factor. The hypothesis of a more rapid differential loss for the interference than for the content is required to explain the increment. Not so apparent, however, is the unstated assumption that the time of opinion retest must coincide with the full or near maximum removal of the inhibiting influence in order to show empirically the increased effect of the communication after an interval of time. With the full removal of the interference, the continued forgetting of content would depress the retained effectiveness of the communication. Consequently, if the second after-questionnaire were given a while after the optimum point had passed, the retained opinion-change evidenced then could be less than that shown initially.

In the Hovland and Weiss experiment, the discounting factor was bound up with the reading of the communication. In the current one, this factor was introduced after the learning was completed and was imparted by a very brief countercommunication. It is not unlikely that, under the latter procedure, the discounting influence is more weakly associated with the content and consequently more rapidly dissociated than under the former condition. Accordingly, the delayed-opinion retest may have come sometime after complete recovery from the interference, resulting in data of the form reported here.[8]

A broader interpretation of the basic theory can be made which may link the sleeper effect to other phenomena. The primary condition needed for obtaining at a later time the response expected initially to a communication (stimuli) is a conflict between motives to accept (approach) and reject (avoid), with the former somewhat stronger than the latter. Any condition which arouses at the outset this conflict of motives can give rise to a delayed effect with the passage of time. In adults subject to cross-pressures in complex social psychological situations, this may produce initially a delay of decision (7) or instability of decision (3). On a perceptual-judgmental level, it may inhibit the con-

[8] The critical factor in this analysis is the time needed for the full removal of the discounting influence. The explanation depends on the assumption that the temporal interval varies with the strength of the association between the interference and the content. However, the magnitude of the interference must also be considered as a determinant of the temporal interval. For a discussion of this factor as a possible explanation of the difference between the empirically obtained forms of the sleeper effect, see Weiss (16).

vergence toward a norm being established in a group situation (10). The later expression of the response inhibited by the conflict is, however, dependent on the strengthening of the acceptance motive over the rejecting one. This may be obtained by increasing the strength of the former or reducing the influence of the latter. In the experiments on the sleeper effect carried out to date, a passive variation of the second of these methods has been employed. In future research, controlled procedures could be actively employed to shift the balance of motives in one opinion direction or the other. A testable implication of this analysis is that a delayed "boomerang" effect would occur, if consequent to the arousal of a conflict condition the rejecting motive were greatly strengthened.

A definite relation between opinions and retention of the related associations is found. Previous studies have indicated that forgetting of the content proceeds more rapidly when opinions are in conflict with the viewpoint of the communication than when they are in agreement (4, 11, 15). This result is verified by the present investigation; for the instances of favorable opinion before and after show a consistently better retention of the statement-label associations than do those of unfavorable opinion before and after. This difference appears shortly after the communication is given and increases with the passage of time.

The investigations reported in the literature did not deal with the effect on retention of changes in opinion, i.e., when the additional motives, derived from opinions, change as a result of the communication from the learning to the retention situation. Consequently, no distinction was made between the possible significance for retention of the relation between the content and before-opinions, as well as after-opinions. On the basis of the results of the present study, it seems reasonable to propose the following extension of the general theory: The rate of forgetting of the content of a communication is affected by the interacting influence of initial and consequent opinions. The research hypothesis related to this broad statement would contain the following assertion: The retention of learned content is aided by favorable before- or after-opinions and adversely affected by unfavorable after-opinions. Under the conditions of this experiment, with acquisition strongly supported by a motive not derived from opinions, a favorable before-opinion seems to impart a somewhat better "retentiveness" value to the associations than does an unfavorable before-opinion. However, a favorable after-opinion gives a very decided advantage in retention over an unfavorable one; and this advantage is of much more significance than the effect derived from the before-opinion.

SUMMARY

1. A communication on the effects of smoking, in the form of a learning experiment, was presented to three groups of high school students (N=228). The students were taught to associate the fact label of "true" or "false" with each of eight statements. After the learning was completed, one group was given a brief countercommunication discounting the truth value of the learned associations. Another group of *Ss* underwent a new learning sequence in which they were to associate the same statements with the names of colors. The third group received no additional treatment. An opinion questionnaire was administered to all *Ss* at least 10 days before the experimental sessions. Opinion changes were assessed immediately after the communication, three weeks after, and six weeks after. Different subgroups of *Ss* from each of the treatment groups were retested at each of these intervals. A retention trial was also included at each of the after-questionnaire periods, so that the extent of forgetting of the statement-label associations could be determined. A group of control *Ss* (N=114) took the before-questionnaire at the same time as the experimental *Ss*, and a different subgroup was retested at each of the three postexperimental intervals.

2. Acquisition of the associations is found not to differ among the experimental subgroups. Also, the experimental treatments do not produce a general difference in the students' retention of the statement-label combinations. Consequently, the two nondiscounted groups were combined for the remaining analyses.

3. When compared with the before-after changes in opinion evidenced by the control *Ss*, the experimental groups show a significant immediate and retained change in over-all

opinion. The differences in extent of opinion-change between the discounted and nondiscounted subgroups approach significance only for the immediate after-test interval.

4. The loss of the effectiveness of the communication, as determined by an analysis of mean change-scores, is shown to be significantly less for the discounted group than for the nondiscounted one by the six-week interval. An analysis based on the individual opinion items also reveals that by the six-week period, the discounted group shows less of a loss in opinion-change on each of the questions than does the nondiscounted group. Thus, a sleeper effect is demonstrated as a consequent of the variation of experimental treatments.

5. Retention of the statement-label associations is found to be affected differentially by the relation between before- and after-opinions and the viewpoint expressed by the communication. The data reveal the following ascending order of opinion conditions making for better retention: opinions unfavorable to the content before and after; opinions favorable before but unfavorable after; and finally, at the same high relative level of advantage, those favorable after whether initially favorable or unfavorable.

REFERENCES

1. ANNIS, A. D., & MEIER, N. C. The induction of opinion through suggestion by means of "planted content." *J. soc. Psychol.*, 1934, 5, 65–81.
2. BARTLETT, F. C. *Remembering.* Cambridge, England: Cambridge Univer. Press, 1932.
3. DINERMAN, HELEN. 1948 votes in the making—a preview. *Publ. Opin. Quart.*, 1948, 12, 585–598.
4. EDWARDS, A. L. Political frames of reference as a factor influencing recognition. *J. abnorm. soc. Psychol.*, 1941, 36, 34–51.
5. HOVLAND, C. I., LUMSDAINE, A. A., & SHEFFIELD, F. D. *Experiments on mass communication.* Princeton: Princeton Univer. Press, 1949.
6. HOVLAND, C. I., & WEISS, W. The influence of source credibility on communication effectiveness. *Publ. Opin. Quart.*, 1951–52, 15, 635–650.
7. LAZARSFELD, P. F., BERELSON, B., & GAUDET, HAZEL. *The people's choice.* New York: Duell, Sloan and Pearce, 1944.
8. LEVINE, J. M., & MURPHY, G. The learning and forgetting of controversial material. *J. abnorm. soc. Psychol.*, 1943, 38, 507–517.
9. PETERSON, RUTH C., & THURSTONE, L. L. *Motion pictures and the social attitudes of children.* New York: Macmillan, 1933.
10. SCHONBAR, ROSALEA A. The interaction of observer-pairs in judging visual extent and movement: the formation of social norms in "structured" situations. *Arch. Psychol.*, 1945, 41, No. 299.
11. SEELEMAN, VIRGINIA. The influence of attitude upon the remembering of pictorial material. *Arch. Psychol.*, 1940, 36, No. 258.
12. SIMS, V. M. Factors influencing attitude toward the TVA. *J. abnorm. soc. Psychol.*, 1938, 33, 34–56.
13. SMITH, F. T. An experiment in modifying attitudes toward the Negro. *Teach. Coll. Contr. Educ.*, 1943, No. 887.
14. WALLEN, R. Ego involvement as a determinant of selective forgetting. *J. abnorm. soc. Psychol.*, 1942, 37, 20–39.
15. WATSON, W. S., & HARTMANN, G. W. The rigidity of a basic attitudinal frame. *J. abnorm. soc. Psychol.*, 1939, 34, 314–335.
16. WEISS, W. A study of a "sleeper" effect in opinion change. Unpublished doctor's dissertation, Yale Univer., 1952.

Received April 8, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

SOCIAL PERCEPTIONS OF AUTHORITARIANS AND NONAUTHORITARIANS

ALVIN SCODEL AND PAUL MUSSEN
The Ohio State University

RECENT social psychological studies have emphasized the relationships between personality structure and social attitudes. The most extensive of these studies, the California study of the *Authoritarian Personality* (1), delineated the differences in personality between individuals high and low in authoritarianism as measured by specially constructed attitude scales. Authoritarian personalities have been described as rigid, extraceptive, repressed, conforming, stereotypical in their thinking, and intolerant of ambiguity; whereas liberal, nonauthoritarian individuals have been described as flexible, intraceptive, and having greater capacity for intense interpersonal relationships.

The California findings were based on responses to questionnaires, clinical interviews, and projective-test materials. Interpersonal relationships of the subjects (*Ss*) were never directly observed. Yet it seems logical to believe that the understanding of the authoritarian individual's dynamics would be broadened if there were further data concerning his actual social behavior.

On the basis of the descriptions of the authoritarian and nonauthoritarian personalities certain predictions about social behavior can be made. Thus it would seem that the authoritarian individual would not be keenly aware of the emotional and social attitudes of others with whom he has brief and casual contact. Since he cannot tolerate ambiguity, the authoritarian individual tends to think of people in rigidly stereotyped categories such as ingroup or outgroup membership. His own security is dependent on his ingroup status and loss of such status is threatening. His perception of others is restricted because of his selective inattention to those behaviors which might set him apart from the ingroup. His approach to others is conventional, nonpersonalized, and essentially lacking in insight. In a social situation with a member of a peer group he maintains his security by assuming that the values and attitudes of others are similar to his own.

The nonauthoritarian personality, on the other hand, has less need to perceive similarity between his own and others' values and attitudes. His reactions to others will be more highly individualized and his evaluations more objective. Therefore, he will be better equipped to use subtle personality cues and make more personalized, insightful assessments of the attitudes and values of others.

The present study is designed to investigate the accuracy with which authoritarian and nonauthoritarian individuals are able to estimate each other's real attitudes. Our hypothesis is that, after a period of social interaction between an authoritarian and a nonauthoritarian person, the authoritarian will perceive the other person as being like himself in social and emotional attitudes, whereas the nonauthoritarian individual will perceive the other person more accurately.[1] Specifically, the authoritarian individual will inaccurately judge a nonauthoritanian to be authoritarian, while the nonauthoritarian will judge the authoritarian accurately, i.e., he will judge him to be authoritarian.

PROCEDURE

Authoritarian and nonauthoritarian populations were selected from a group of 276 students in the introductory psychology course at the Ohio State University who had responded to a questionnaire which was presented as "a survey of students' feelings about themselves and important topics." Actually the questionnaire consisted of a randomly ordered presentation of 28 items from the F scale, a scale published in the *Authoritarian Personality* (1) and designed to measure "potential for fascism" or authoritarianism, and 30 items from the Minnesota Multiphasic Personality Inventory (MMPI). The majority of the MMPI items involved recognition of one's own psychological inadequacies, hostility toward parents, and general

[1] It should be clearly noted that the present hypothesis applies only to members of the majority group. Authoritarian personalities would undoubtedly judge minority group members to have attitudes different from their own.

interests. Although *Ss* were told that they were answering the questionnaire anonymously, they could be identified by date of birth, religion, and major field, all of which were entered on the questionnaire. The *Ss* were instructed to indicate their degree of agreement or disagreement with each item (from +3 to −3).

Of all those answering the questionnaire, the 40 students who received the highest F scores (the authoritarian or high group) and the 40 who received the lowest F scores (the nonauthoritarian or low group) were selected as *Ss* for the experiment. The cut-off points for the high and low groups were 119 and 79, respectively, points equidistant from the mean for the total population of 276, which was 99.06. Because of scheduling difficulties and the elimination of *Ss* who belonged to minority groups, the actual sample was reduced to 27 high and 27 low *Ss*, 12 men and 15 women in each group.

The experimental situation was a social one in which a high *S* was paired with a low *S* of the same sex and the pair was instructed to carry on a 20-minute discussion of radio, television, and movies. In order to disguise the purpose of the experiment, *Ss* were told that they were selected to participate in an experiment investigating the nature of social interaction between people meeting for the first time. The experimenter (*E*) remained in the room while the conversation was in progress to insure that the *Ss* did not digress from the assigned topics. All *Ss* complied with the instructions and almost without exception were able to maintain a running conversation without sustained pauses.

After the termination of the conversation, each *S* was taken to a separate room and again given the questionnaire. This time, however, he was instructed to play a role and to respond to the questionnaire *as he felt the other subject with whom he had just been speaking would respond to it.* There were no indications that the *Ss* knew that the second administration of the questionnaire was related to the first administration.

Thus, at the end of the experiment, each *S* had responded to the same questionnaire twice. The first questionnaire contained his own responses to the items, and from it his actual F score and the responses to the MMPI items were determined. Completing the second questionnaire involved a kind of role playing, and from it each *S*'s judgments of how his partner would respond to the questionnaire were obtained. Actual and estimated (attributed-to-the-partner) F scale scores were obtained by adding a constant of 4 to the value which had been assigned to each F item, e.g., a −1 response to an item was scored 3, a +2 response was scored 6.

RESULTS

In order to test our hypotheses with respect to the perception of social attitudes, lows' estimates of highs' responses to the F scale were compared with highs' estimates of lows' responses to the F items. If the hypothesis is to be substantiated, highs should judge lows inaccurately, i.e., they should predict that they would be high in F; while lows, being more accurate in their predictions, should judge the highs to be high.

Two methods of analyzing the data were employed. Table 1 presents the results of the first method of analysis which compares the *actual* mean scores of the high and low groups with the mean scores attributed to them by their partners in the experiment.

TABLE 1

MEANS AND SIGMAS OF ACTUAL AND ESTIMATED F SCORES FOR THE HIGH AND LOW GROUPS

GROUP	ACTUAL MEAN	ACTUAL SIGMA	EST.* MEAN	EST.* SIGMA	*diff*	*t*	*p*
Highs (N=27)	129.26	9.34	124.19	11.08	5.07	1.79	.10
Lows (N=27)	66.52	11.15	99.81	21.88	33.29	7.05	<.01

* Score attributed to the group by their partners.

Table 1 makes it clear that the difference between the highs' actual mean score and their mean estimate of the lows' score is not statistically significant. In short, the highs do not perceive the lows as having attitudes significantly different from their own. The lows, on the other hand, estimate the highs' mean score to be significantly higher than their own, but also sifinificantly lower than the highs' actual mean score.[2] This analysis of the data may be said to give limited support to our hypothesis: the highs inaccurately perceived the lows as having high scores. The lows accurately perceived the highs as being higher than themselves, but did not see them as being as high as they actually were. The

[2] The difference between the highs' actual mean score and the mean score attributed to them by the lows yielded a *t* value of 6.44, significant beyond the .01 level.

mean score attributed to the highs by the lows was 99.81, which approximated the mean score for the reference group of 276 subjects.

In the second analysis of the data, each estimated score was classified into one of the three following categories: (*a*) *accurate* (estimated score was in the same category as the actual score, i.e., the low's estimate of the high's responses was such that the high's total score was 119 or over, and the high's judgments of the low's responses gave the low a total F score of 79 or below); (*b*) *middle* (the estimated score fell in the middle range, 80 to 118); (*c*) *inaccurate* (in his role taking, the low's estimates of the high gave the high a total score of 79 or below and thus placed him in the low category, and, conversely, the high's estimates of the low placed the low in the high category). Table 2 shows the distributions of *accurate, inaccurate,* and *middle* estimates made by the highs and lows.

TABLE 2

NUMBER OF ACCURATE, MIDDLE, AND INACCURATE JUDGMENTS OF F SCORES MADE BY HIGHS AND LOWS.

GROUP	ACCURATE	MIDDLE	INACCURATE
High	0	1	26
Low	9	13	5

The chi-square test applied to this table yields a chi-square value of 33.5, which is significant beyond the .01 level and indicates the extremely large proportion of inaccurate estimates on the part of the high group. In a sense, however, the table gives only a partial confirmation of the hypothesis since, although 26 of the 27 highs predicted inaccurately, only 9 of the lows predicted accurately. On the other hand, only one of the highs estimated a low to be in a lower category than he was himself, whereas 22 of the 27 lows estimated highs to be in a higher category than they were themselves.

Responses to the MMPI items were analyzed to determine whether the lows could assess the responses of their high partners to each item more accurately than the highs could assess those of their low partners. If our hypothesis is valid, lows should, in fact, make more accurate appraisals of the responses of the highs than highs make of the responses of the lows. The estimate for an item was considered to be accurate when the sign preceding the estimate was the same as the sign preceding the actual response, e.g., an estimate of —1 was considered to be in agreement with an actual response of —3.

Chi-square analyses revealed that the lows predicted the responses of the highs significantly better than the highs predicted the responses of the lows on the following four items:

My hardest battles are with myself.
I am certainly lacking in self confidence.
I like poetry.
Most people make friends because friends are likely to be useful to them.

In the case of only one item (It makes me uncomfortable to put on a stunt at a party even when others are doing the same sort of things.) were the highs more accurate estimaters than the lows. In 25 of the 30 items there were no significant differences between the prediction accuracies of the two groups. Although the number of discriminating items is too small to permit any definitive statements, there seems to be a trend in a direction consistent with our hypothesis.

DISCUSSION

The over-all results generally support the major hypothesis of this study: nonauthoritarian individuals do, in fact, make more accurate judgments about authoritarians than authoritarians do about nonauthoritarians. The study also yields further data substantiating some of the other findings of the investigation of the *Authoritarian Personality* (1). The present findings, like those of the California investigation, indicate that, as groups, the authoritarians are much more homogeneous than the nonauthoritarians, who show a great deal of variability. Thus the high group shows almost complete unanimity in the inaccurate assessments of their partners, whereas the judgments of the low group are distributed over the accurate, middle, and inaccurate categories of judgment.

The present investigation gives additional information on the restricted nature of the authoritarian individual's social perceptions. It has been pointed out that these individuals tend to project their own hostile and other

repressed, unacceptable impulses onto minority groups—a kind of behavior which Cameron (2) has called "disowning projection."

The present data seem to indicate another and equally important aspect of their social behavior: the attribution of their own opinions, attitudes, and characteristics to other members of the ingroup. Cameron has labeled this mechanism "assimilative projection."

Both these mechanisms may be interpreted in the light of the highs' inability to tolerate ambiguity. Ambiguity, and thus anxiety, are reduced by a rigid dichotomization, which minimizes individual differences and evaluates people only in terms of their ingroup or outgroup memberships. The highs' strong need to identify with an ingroup of superior status results in perceptual distortions which exaggerate and extend the similarities between themselves and their peers. On the other hand, their strong need to deny their hostile tendencies results in another kind of perceptual distortion, the exaggeration of differences between themselves and members of minority groups with whom they feel no identification at all.

It is felt that the present study demonstrates how hypotheses about interpersonal behavior may be derived from clinically obtained information and how the testing of these hypotheses may broaden our understanding of the interrelationships between personality and social behavior. A valuable extension of the present study would employ a similar experimental situation in which lows would be paired with lows and highs would be paired with highs. If our hypothesis is correct, the judgments yielded in these situations would be accurate, i.e., lows would judge lows to be low and highs would judge highs to be high. This would be the only instance in which accurate estimates by the highs would be predicted.

SUMMARY

This study was designed to test the hypothesis that authoritarians, because of their lack of insight into others and their need to consider themselves members of the ingroup, would perceive nonauthoritarian peers to have attitudes and personality characteristics similar to their own. Nonauthoritarians, on the other hand, being more insightful and more individualized in their approaches to others, would judge authoritarians accurately with respect to attitudes and personality characteristics.

In order to test this hypothesis, 27 pairs of *Ss*, each consisting of an authoritarian and a nonauthoritarian, were placed in a social situation and instructed to discuss the neutral topics of radio, television, and movies. Each *S* had previously responded to a questionnaire consisting of the F scale and 30 items from the MMPI. After the discussion was completed, each *S* was again given the questionnaire, but this time he was instructed to respond to it as he felt the other *S* with whom he had just been speaking would respond to it. Thus we were able to compare: (*a*) lows' (nonauthoritarians) *actual* responses and F scores and their *estimates* of highs' (authoritarians) responses and F scores; (*b*) highs' *actual* responses and F scores and their *estimates* of lows' responses and F scores.

Analyses of the data revealed:

1. Highs did not perceive lows as having F scores significantly different from their own.

2. Lows estimated highs' F scores to be significantly higher than their own, but lower than the highs' actual scores.

3. When F score estimates were classified as accurate, middle, or inaccurate, 26 of the 27 highs made inaccurate estimates of the lows, while only five lows made inaccurate estimates of the highs.

4. In the cases of four of the 30 MMPI items, lows made more accurate assessments of the responses of highs than highs made of the responses of lows. In the case of only one item was the situation reversed.

The over-all results generally support the hypothesis that nonauthoritarians make more accurate judgments about authoritarians than authoritarians do about nonauthoritarians.

REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
2. CAMERON, N. *The psychology of behavior disorders*. Boston: Houghton Mifflin, 1947.

Received April 15, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

THE ATTITUDES OF THE MOTHERS OF MALE SCHIZOPHRENICS TOWARD CHILD BEHAVIOR [1, 2]

JOSEPH C. MARK

Bronx VA Hospital and New York University

THIS investigation was designed to determine if the attitudes of the mothers of schizophrenics toward child behavior differ from the attitudes of the mothers of nonschizophrenics.

The male schizophrenics, whose mothers are the subject of this study, are veterans of World War II, hospitalized at VA installations at Northport, Long Island, and the Bronx, hospitals which serve the New York area.[3]

METHOD

The results are based on material obtained through an attitude survey of a "disguised-structured" type (2). This attitude survey contains 139 items pertaining to child rearing. The items are worded in the form of stereotypes, such as "A child should be seen and not heard." Shoben (10) found that such an instrument was disarming because of the disguised nature of the items, while allowing for a projection of opinion. At the same time the test is structured and provides for responses on an objective basis. Responses to each item are made on a four-point scale: strongly agree (A), mildly agree (a), mildly disagree (d), strongly disagree (D).

Subjects

The attitude survey was administered to 100 mothers of male schizophrenics and to a control group of 100 mothers of male nonschizophrenics.

The control group was selected from among the mothers visiting general medical patients at the VA Hospital in the Bronx, women's auxiliaries of veterans' organizations, volunteer groups, and from private contacts. None of the sons of the control *Ss* had a history of hospitalization for mental illness, according to the mothers. The control group is comparable to the experimental group with respect to age, religion, education, socioeconomic status, and age of sons.

The mean age of the mothers in the experimental group is 55.7 years, and that of the mothers in the control group is 54.9 years.

The experimental and the control groups each have 35 Catholic, 35 Jewish, and 30 Protestant *Ss*.

The mothers in both groups are predominantly of grammar school educational background (50 experimental and 47 control) with the bulk of the remaining members in both groups at the high school level (37 experimental and 43 control). Two in each group went to college; twelve had no formal education (7 experimentals and 5 controls); seven were "unreported" (4 experimentals and 3 controls).

As an index of socioeconomic status, occupation of husband was employed. There were 11 professionals in each group, 19 proprietors and managers in the experimental group and 26 in the control group; 12 clerks and kindred workers in the experimentals and 11 in the controls; 30 skilled workers in the experimentals and 28 in the controls; with the rest divided fairly evenly between skilled and unskilled workers.

The mean age of the schizophrenic sons was 29 years and the sons of the controls averaged 28 years. The mean numbers of children in both groups approximated three children per family.

RELATED STUDIES

The influence of parents on a child's mental health, originally stressed by Freud (3, pp. 80–81) and recently emphasized by Sullivan (12), is now widely accepted. Speaking from his experience as a psychiatrist during World War II, Strecker (11), for example, found that the mental health of many soldiers was jeopardized by the attitude of his "Mom."

In a series of 45 unselected case histories of schizophrenics, Kasanin, Knight, and Sage reported (7) that maternal overprotection or rejection was present in 60 per cent of the cases (rejection was present in only two cases).

Studying the personalities of 25 mothers

[1] Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the author are the result of his own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

[2] This article is based on a doctoral dissertation submitted to the School of Education, New York University. Grateful acknowledgments are due to the Chairman of the sponsoring committee, Professor Brian E. Tomlinson, and to the other committee members, Professors Philip J. Zlatchin and Bernard N. Kalinkowitz.

[3] The author wishes to acknowledge gratefully the cooperation of Dr. H. L. Flowers, Chief, Neuropsychiatric Service, and Dr. R. S. Morrow, Chief Clinical Psychologist, Bronx VA Hospital; and Dr. Earl S. Swartzlander, formerly Chief Clinical Psychologist at VA Hospital, Northport, Long Island.

through the method of repeated interview, Tietze (13) concluded that the mothers were generally overanxious, obsessive, and domineering. She felt that the subtly dominating mother was particularly dangerous to the child, because the child was left with no opportunity to express aggressive feelings openly.

Gerard and Siegel (5) examined the family background of schizophrenics through anamneses collected from 71 parents and relatives. They found the parents to be immature with a poor marital relationship; and that the child was often exposed to markedly overprotective attitudes.

In contrast to Gerard and Siegel, Prout and White (8) found little difference between the life histories of 25 mothers of male schizophrenics and the life histories of the mothers of 25 nonschizophrenic males. They concluded from Rorschach data, however, that the mothers of the schizophrenics tended to compensate for their own life disappointments by living through their sons rather than having an independent life of their own.

There is then evidence for a pathological familial influence in the lives of schizophrenics. Of course, those who adhere to an organic or hereditary explanation of schizophrenia can maintain that a constitutional weakness in the child makes itself apparent in the form of a sickly child. The overprotective attitudes often attributed to the mothers of schizophrenics are then interpreted as a reaction to a basic inferiority in the child (6). On the other hand, those who have a psychogenic bias can insist that the mother's attitudes are projected onto the child because of her own unconscious needs, regardless of the physical and psychological makeup of the child (4, 9). Fromm-Reichmann (4) has coined the term "schizophrenogenic mother" to describe the mother who has such an injurious effect on her child as to cause him to become schizophrenic; and Rosen (9) has sought the origins of the schizophrenic pattern in the relations between mother and child, especially searching for those feelings of the mother which convey rejection.

Without attempting to confront these crucial questions concerning the etiology of schizophrenia, it is the purpose in this study to ascertain in some measure whether or not a characteristic set of attitudes concerning the rearing of children can be attributed to the mothers of schizophrenics.

RESULTS

The two groups' responses to the 139 items of the attitude survey were compared, item by item, by means of the chi-square model. Of the 139 items, 67 proved significant at the .05 level of confidence or better. The 139 items, grouped by significance level, are presented in Table 1. In *all* instances of significant items, the mothers of schizophrenics more frequently *agreed* with the item than did the mothers of nonschizophrenics.

To determine if the differences in attitudes assume identifiable patterns, the 67 significant items were classified by four judges (clinical psychologists) into three preconceived clusters. These clusters were obtained from the Fels Institute's longitudinal research appraising parental behavior (1, pp. 4–5). These three clusters are as follows:

1. The measures of control the parent employs. This continuum extends from the parent who is restrictive and coercive, allowing the child little freedom of choice or range of activity, to the parent who is lax, ineffectual, unable to control the child. Between these two extremes is the parent who allows the child freedom but who is capable of asserting authority when and where desired.

2. The intellectual objectivity of the parent. The objective parent is rational rather than emotional in his behavior. Regardless of his acceptance or hostility, his overt behavior is largely governed by what he consciously decides is an appropriate policy. The emotional parent, in contrast, is unable to divorce his behavior from his immediate mood and consequently directly expresses his emotional attitude in his behavior, be it friendly or antagonistic.

3. The warmth of the parent-child relationship. A warm home is one in which the parent genuinely likes and enjoys the child, finds contact with him rewarding and pleasant, is appreciative and approving of the child's personality—in short, is "acceptant" in the conventional sense of the term. The position of the home varies on this continuum from excessive devotion through casual enjoyment and cool detachment down through vigorous hostility.

Some of the 67 significant items were found to apply to more than one of these categories: 16 items overlapped; 7 items remained unclassified. The following is a tabulation

TABLE 1

ITEMS SIGNIFICANTLY DIFFERENTIATING MOTHERS OF SCHIZOPHRENICS FROM MOTHERS OF NONSCHIZOPHRENICS

.001 LEVEL

13. Children should be taken to and from school until the age of eight just to make sure there are no accidents.
18. A mother should make it her business to know everything her children are thinking.
20. If children are quiet for a little while a mother should immediately find out what they are thinking about.
32. Children should not annoy parents with their unimportant problems.
33. A devoted mother has no time for social life.
35. A watchful mother can keep her child out of all accidents.
40. Playing too much with a child will spoil him.
43. A parent must never make mistakes in front of the child.
51. Parents should sacrifice everything for their children.
54. When the father punishes a child for no good reason the mother should take the child's side.
64. A mother has to suffer much and say little.
81. Most children are toilet trained by 15 months of age.
101. Children who take part in sex play become sex criminals when they grow up.
115. A child should not plan to enter any occupation that his parents don't approve of.
121. Too much affection will make a child a "softie."

.01 LEVEL

19. Answering questions about sex is embarrassing and unnecessary.
22. A son is usually like the mother and a daughter like the father.
23. Punishing a child is a father's job.
29. It is all right for a mother to sleep with a child because it gives him a feeling of being loved and wanted.
30. A good way to get children to obey is by giving them presents or promising them treats.
47. Spanking a child does more good than harm.
57. Most children are able to talk by the age of 12 months.
58. There's little thanks or pleasure in raising children.
61. Parents ought to close their eyes to their children's faults.
62. One reason that it's sad to see children grow up is because they need you more when they are babies.
66. Children should not be punished for doing anything they've seen their parents do.
67. Boys should never be allowed to play with dolls and carriages.
71. The sooner a child learns to walk the better he's been trained.
75. It is not the duty of the parent to teach the child about sex.
111. Some children are just naturally bad.
112. Babies are more fun than older children.
117. A good child will never allow his parents to go to a home when they grow too old.
118. A good mother should shelter her child from life's little difficulties.
123. A child should never keep a secret from his parents.
131. Children seven years of age are really too young to spend summers away from home in a camp.
135. Mothers should have vacations away from their children.

.02 LEVEL

14. After punishing a child a parent naturally wants to make up for it by giving him everything he wants.
27. A mother should never be separated from her child.
37. Every parent has a favorite child.
41. Children need some of the natural meanness taken out of them.
50. A mother should shower her child with praise at all times.
53. It's better to have a boy than a girl.
73. Most children should have more discipline than they get.
90. It is better for children to play at home than to visit other children.
96. A child will develop a better character if he works after school instead of playing.
108. If children are to grow up and get somewhere in life they must be continuously kept after.
110. A young child should be protected from hearing about sex.
116. A child should not be allowed to play in the living room.
125. When an infant doesn't like a certain food the mother should stop feeding it to him.

.05 LEVEL

4. When a child won't eat you should tell him how nicely other children eat.
9. Children will neglect their schoolwork if parents don't keep after them.
10. Quiet children are much nicer than little chatterboxes.
12. The least a child can do for his parents when he grows up is to take care of them in their old age.
25. A good child doesn't fight with other children.
45. Sex is one of the greatest problems a parent has with children.
69. Children should do nothing without the consent of their parents.
74. Whenever a child is slow in dressing the parent should do it for him.
80. A child's friends usually do more harm than good.
104. No person should get married before the age of 25.
106. It is best to give children the idea that their parents have no faults.
107. A child shouldn't be allowed to see his parents completely undressed.
127. It is wicked for children to disobey their parents.
130. A child should be weaned away from the bottle or breast as soon as possible.
132. Strict discipline develops a fine strong character.
134. Children should always be punished for being bad.
138. Children who always obey grow up to be the best adults.
139. Too much freedom will make a child wild.

of the number of items in each of the three clusters employed:

1. Measures of Control (39 items) [4]
   a. lax and ineffectual, 6
   b. restrictive and coercive, 33
2. Intellectual Objectivity (14 items) [5]
   a. rational, 8
   b. emotional, 6
3. Warmth of the Relationship (25 items) [6]
   a. excessive devotion, 13
   b. cool detachment, 10
   c. vigorous hostility, 2
4. Unclassified (7 items)

Judging from their numerical distribution, the mothers of schizophrenics as a group can be characterized as being mainly restrictive in their attitudes as to methods of controlling children. Lax and ineffectual attitudes are also present, but to a much lesser extent, and are outnumbered by the restrictive items 33 to 6.[7]

The mother of the schizophrenic believes in allowing a child no freedom or choice of activity and in prescribing and carefully channeling his behavior (items 13, 35, 69, 74). She frowns on friends and sex play and believes in keeping sexual information from him (items 80, 90, 101, 110). She feels it is a mother's duty to know everything that the child is thinking (items 18, 20, 123). This pervasively restrictive attitude extends into adulthood (items 12, 104, 108). Although she believes in stricter discipline, it is only for purposes of restricting behavior; lax methods are evidently approved if they gain desired ends (items 29, 30, 61).

The cluster of *intellectual objectivity* is composed of a rational or deliberate approach, on the one hand, and an emotional or impulsive one, on the other. The judges rated eight and six items, respectively, into these categories. This cluster was difficult to define for our purposes and the judges had trouble in rating. The few items classified into this cluster reflect this difficulty and it is not possible to make any definite conclusions.

The *warmth of the relationship* consists of the following continuum: excessive devotion, casual enjoyment, cool detachment, and vigorous hostility. The judges placed 25 items into the warmth cluster and identified 13 as indicating excessive devotion, none as casual enjoyment, 10 as cool detachment, and 2 as vigorous hostility. If casual enjoyment is taken as a center point, there are 13 items to one side of it and 12 to the other.

These mothers as a group thought that "a devoted mother has no time for social life" (item 33), but that "children should not annoy their parents with their unimportant problems" (item 32). They thought that "a mother should shower her child with praise at all times" (item 50), but that it was not desirable to play too much with a child as it would spoil him (item 40). They felt that "it's sad to see children grow up because they need you more when they're babies" (item 62), and that "quiet children are much nicer than little chatterboxes" (item 10).

The presence of such opposing attitudes as devotion and detachment may be pointing up the ambivalence of feeling toward their own children.

As far as the design of the study is concerned, a cross validation was not attempted. A replication would be desirable, however, to see how many of the significant items would repeat themselves. It is possible, for example, that about 7 of the 67 items at the .05 level or better may be significant by chance alone. The author has not been able to obtain a suitable population for such a cross validation but welcomes the efforts of other psychologists who would desire to do such a replication.[8]

SUMMARY

1. The attitudes of the mothers of schizophrenics were found to differ from the atti-

[4] The numbers of these items are as follows: lax and ineffectual: 23, 29, 30, 47, 61, 125; restrictive and coercive: 4, 9, 12, 13, 18, 20, 25, 35, 41, 67, 69, 71, 73, 74, 80, 90, 96, 101, 104, 107, 108, 110, 115, 116, 118, 123, 127, 130, 131, 132, 134, 138, 139.

[5] Item numbers under "rational": 40, 43, 57, 71, 81, 96, 121, 138; under "emotional": 14, 29, 54, 62, 65, 126.

[6] "Excessive devotion" item numbers: 13, 27, 29, 33, 35, 50, 51, 54, 62, 112, 118, 131, 135; "cool detachment": 10, 19, 32, 40, 75, 96, 111, 121, 130, 132; "vigorous hostility": 41, 58.

[7] A middle point on the Fels Behavior Rating Scale exists between these two extremes and is defined as a parent who "allows sufficient freedom but is capable of asserting authority whenever it is deemed necessary." No items were classified into the middle position by the judges.

[8] For the complete questionnaire used in the study and the table of responses of the mothers of schizophrenics and the mothers of nonschizophrenics, order Document 3672 from the American Documentation Institute, % Library of Congress, Washington 25, D.C., remitting $1.00 for microfilm (images 1 inch high on standard 35-mm. motion picture film) or $1.50 for photocopies (6 x 8 inches) readable without optical aid.

tudes of the mothers of nonschizophrenics with respect to various child-rearing practices.

2. Sixty-seven of the 139 items in an attitude survey differentiated the two groups at the .05 level of confidence or better.

3. The mothers of schizophrenics were found to be mainly restrictive in their control of the child. When it came to the warmth of the relationship, they exhibited attitudes of both excessive devotion and cool detachment.

REFERENCES

1. BALDWIN, A. L., KALHORN, JOAN, & BREESE, FAY H. The appraisal of parental behavior. *Psychol. Monogr.*, 1949, 63, No. 4 (Whole No. 299).
2. CAMPBELL, D. T. The indirect assessment of social attitudes. *Psychol. Bull.*, 1950, 47, 15–38.
3. FREUD, S. *The interpretation of dreams.* New York: Macmillan, 1937.
4. FROMM-REICHMANN, FRIEDA. Notes on development of treatment of schizophrenics by psychoanalytic therapy. *Psychiatry* 1948, 11, 263.
5. GERARD, D. L., & SIEGEL, J. The family background of schizophrenia. *Psychiat. Quart.*, 1950, 24, 47–73.
6. KALLMAN, F. S. Genetic theory of schizophrenia. *Amer. J. Psychiat.*, 1946, 103, 309–322.
7. KASANIN, J., KNIGHT, E., & SAGE, P. Parent-child relationship in schizophrenia. *J. nerv. ment. Dis.*, 1934, 79, 249–263.
8. PROUT, C. T., & WHITE, MARY A. A controlled study of personality relationships in mothers of schizophrenic male patients. *Amer J. Psychiat.*, 1950, 107, 251–256.
9. ROSEN, J. N. The treatment of schizophrenic psychoses by direct analytic therapy. *Psychiat. Quart.*, 1947, 21, 3–37.
10. SHOBEN, E. J., JR. The assessment of parental attitudes in relation to child adjustment. *Genet. Psychol. Monogr.*, 1949, 39, 101–148.
11. STRECKER, E. A. *Their mothers' sons.* New York: Lippincott, 1946.
12. SULLIVAN, H. S. *Conceptions of modern psychiatry.* Washington, D. C.: William Alanson White Psychiatric Foundation, 1947.
13. TIETZE, TRUDE. A study of mothers of schizophrenic patients. *Psychiatry*, 1949, 12, 55–65.

Received April 16, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

PERSONALITY DIFFERENCES BETWEEN PATIENTS CLASSIFIED AS PSYCHOSOMATIC AND AS NONPSYCHOSOMATIC [1]

LEONARD KRASNER

Veterans Administration Regional Office
Brooklyn, New York

In recent years there has been a great increase in articles and books concerning the relationship between certain physical disorders and the personality and background of the patients suffering from these illnesses (2, 9, 10, 41). These disorders have been generally given the title of "psychosomatic" implying some sort of relationship between the organic illness and some psychological processes.

In this study we wish to determine by means of objective psychological tests whether there are measurable personality differences between two groups of patients suffering from diseases usually termed psychosomatic, viz., ulcerative colitis and duodenal ulcer; and also between each of these groups and a third group of patients free of psychosomatic disorders.

Although the concept of psychosomatic medicine is of comparatively recent origin, it already includes a large number of conditions hitherto considered to be primarily somatic in origin. Weiss and English (41), authors of one of the standard texts on psychosomatic medicine, believe three types of patients fall into the psychosomatic category. The first group includes those patients who have bodily complaints, but do not have any definite bodily disease to account for their complaints. These are the so-called "functional" disorders. A second group is made up of patients whose symptoms are in part dependent upon emotional factors, even though organic findings are present, e.g., organic heart disease. The third category includes those diseases, such as asthma, hypertension, peptic ulcer, and ulcerative colitis, in whose etiology psychological factors are considered to be of primary, if not crucial, importance. It is from this last category that the psychosomatic patients in this study have been selected.

Dunbar (10), Alexander (2), and many of the other writers in this field state that specific personality patterns are associated with each of the disease categories which they term psychosomatic, e.g., hypertension, asthma, peptic ulcer, ulcerative colitis, and various cardiovascular disorders. They further state that the personalities of patients in each of these disease groups differ in certain respects from patients who have different psychosomatic diseases. Furthermore, they imply that a person develops a particular type of psychosomatic disease because there is something about his personality and/or background which predisposes him to that particular disease rather than to some other. The literature abounds with descriptive adjectives, often of a very vague and undefined nature, of the types of personality found in each of these various diseases. However, there have been only a few attempts to demonstrate by means of psychological tests whether there actually exist definite personality traits associated with a given disease category which marks this disease off from other categories. Kornreich and Krasner in a forthcoming article will summarize the results of these studies and their relevance to various hypotheses as to the personality concomitants associated with these disorders.

As stated above, the purpose of this study is to determine measurable personality differences, if they exist, between two groups of patients hospitalized with disease entities termed psychosomatic. As a control we have also determined whether there are personality differences between either of these two psychosomatic groups and a group of hospitalized patients who do not at present exhibit (and who have never previously exhibited) diseases classified as psychosomatic. The per-

[1] The writer wishes to acknowledge with gratitude the helpful suggestions of Dr. Henry E. Garrett and Dr. Joseph Zubin in formulating and conducting this study, which is part of a Ph.D. dissertation at Columbia University, Graduate Faculties. He is also indebted to Dr. Bernard Locke and Dr. Melvin Kornreich for their assistance.

sonality differences which we shall investigate are those which appear in the reponses of these three groups of patients to an objectively scored personality inventory, the Guilford-Martin Factor Inventory. Personality differences will be considered to exist when the differences in the responses between the groups are statistically significant.

It may be added that this study is oriented toward ascertaining personality differences which characterize the present morbid condition of the patients. Such differences may or may not have been characteristic of these groups prior to their present illness.

METHOD

Subjects

Our subjects (*Ss*) were three groups of patients hospitalized at Halloran VA Hospital. These *Ss* were white males, veterans, native-born, between the ages of 20 and 40. All had completed at least the eighth grade. This last requirement was decided upon to insure an adequate reading ability of the inventories and questionnaires used in this study.

Our first group consisted of 30 patients with duodenal ulcer as their only medical diagnosis. Patients who were diagnosed as having some other disease in addition to duodenal ulcer were not included in the study in order to insure as uncomplicated a group of ulcer patients as possible. The exact cause of duodenal ulcer is not known despite many theories as to etiology, including those of a psychogenic origin (1, 4, 26, 31, 38). Recent published studies containing psychological test results (mainly Rorschachs) on ulcer patients have been those of Ruesch (38), Modell and Potter (32), Osborne and Sanders (36), and Poser (37).

Our second group consisted of 27 patients with the medical diagnosis of ulcerative colitis. The patients in this group were free from any other major illness. As with duodenal ulcer, the true cause of this illness is unknown despite many and varied theories (3, 4, 6, 7, 14, 24, 39, 43). Mahoney (30) has recently published an article giving Rorschach data on ulcerative colitis patients.

The problem of diagnosing duodenal ulcer and ulcerative colitis is a medical one. In this study all of the medical diagnoses were made by the same physician, the chief of the gastrointestinal service.[2] A recent study by Moschowitz and Roudin (33), on the incidence of various psychosomatic diseases in the same individual, clearly indicated that these two particular disease entities are almost always mutually exclusive in the same individual.

The third group was composed of 44 patients with nonpsychosomatic disorders. These men were hospitalized either for inguinal hernia or for pilonidal cyst, both of which are ailments that have not been considered in the literature to be markedly affected by emotional states. Moreover, the patients in this group had no histories of psychiatric or psychosomatic disorder.

[2] We are grateful for the aid rendered by Dr. H. A. Weiner, Chief of Medical Service, and Dr. Milton Redish, Chief of Gastro-intestinal Service, of Halloran Hospital in obtaining and diagnosing the patients in this study.

Procedure

It was explained to the patients tested that the hospital was interested in finding out certain information about the tests and their uses. All were assured that results would be strictly confidential and used for the eventual benefit of the patients in the hospital. The following battery of tests and questionnaires was then administered:

1. The Wechsler-Bellevue Intelligence Scale, Form I (42). This test, clinically the most widely used with adults, enabled us to obtain the general level of intellectual functioning.

2. The Guilford-Martin Factor Inventory (STDCR, GAMIN, OAgCo) (15–23). This particular inventory was chosen rather than some other personality test for several reasons. First, the items on this questionnaire were selected on the basis of a factor analysis. Thus this inventory depends upon internal factorization rather than on external psychiatric categories. Eysenck (12) has pointed out that the statistical work involved in constructing the questionnaire "is of a very high standard." Second, the factors obtained by Guilford are similar to those obtained in independent studies by Cattell (5), Mosier (34), Eysenck (11), and others (44) who have used the factor analytical technique. Finally, the Guilford inventory contains 511 items which lend themselves to an item analysis. The names which Guilford and Martin give to their factors are as follows: S (Social Introversion-Extroversion), T (Thinking Introversion-Extroversion), D (Depression), C (Cycloid Disposition), R (Rhathymia-Impulsiveness), G (General Activity), A (Ascendance-Submission), M (Masculinity-Femininity), I (Inferiority Feelings), N (Nervousness), O (Objectivity), Co (Cooperativeness), Ag (Agreeableness).

Studies by Lovell (29) and North (35), as well as the fairly high correlations between the factors listed by Guilford in his original studies, indicate that the factors are not independent. However, for the purposes of this study, the 13 original factors of Guilford provide definable dimensions in terms of which statistical comparisons can be made.

3. Thurstone Interest Schedule (40). This schedule was used to provide broad classifications of vocational interests in general areas. It requires less time to administer and is easier to score than longer and more complicated inventories. It yields a profile which represents interests in 10 vocational fields: physical science, biological science, computational, business, executive, persuasive, linguistic, humanitarian, artistic, and musical.

4. A questionnaire designed to obtain background information and attitudes in various areas not covered by either the Guilford or Thurstone inventories. This questionnaire contained 66 items,

each with three or more multiple-choice answers, depending on the exact nature of the item. Such areas as schooling, recreational activities, hobbies, smoking and drinking habits, and vocational attitudes were covered. About one-fourth of the items came from a questionnaire devised for use in the Armed Services. The remaining items were drawn from the literature and covered those activities supposedly characteristic of these groups.

5. Since there are some indications in the literature of a connection between psychosomatic illness and morphological or physical data (8), the height and weight of each patient in the study were noted. These data were obtained from the hospital charts.

Group comparisons on the Rorschach and other projective procedures used in the study of these patients are reported elsewhere (27) and in forthcoming articles.

RESULTS

Table 1 contains background information as to age, education, and reported weekly income of the three groups of patients. The religious distribution of the three groups, as well as that of the entire hospital population, irrespective of diagnosis, for the same period of time, is given in Table 2. The Jewish patients showed a greater incidence of colitis than the other two religious groups (p=.01). The Catholic patients showed a greater incidence of duodenal ulcer than the Jewish patients (p=.05).

TABLE 1

AGE, EDUCATION, AND REPORTED AVERAGE WEEKLY INCOME OF THREE GROUPS OF HOSPITALIZED PATIENTS

GROUP	N	AGE	EDUCATION (Years)	INCOME
Colitis	27	27.33	11.91	$58.40
Ulcer	30	28.27	10.83	$60.40
Nonpsychosom.	44	25.45	11.50	$56.74

The colitis group attained a significantly higher IQ on the Wechsler-Bellevue test than either of the other two groups (Table 3). By means of the chi-square test (13) it was found that none of the three distributions of IQ departed significantly from normality. Of the three groups, the IQ's of the colitis group were most nearly normal. (p for colitis group = .70, p for nonpsychosomatic group = .60, p for ulcer group = .20.)

Each of the 13 factors in the Guilford-Martin inventory was scored by a separate scoring key provided by the authors. The significance of the difference between the raw scores made by each of the groups for each factor was obtained by means of the t test (13). The comparisons were: colitis with nonpsychosomatic, ulcer with nonpsychosomatic, ulcer with colitis, and the combined psychosomatic group with the nonpsychosomatic group. The variance ratio F (28) was computed for each comparison. Since the IQ differences between the ulcer and colitis, and the colitis and nonpsychosomatic groups were significant, an adjustment was made to equalize the intelligence factor by the method of covariance analysis (28). Table 4 contains the raw scores, the differences in scores between each of the groups, the results of the t tests, and the variance ratios. The raw scores on each factor for each group were then converted into standard scores (C scores), from 0 to 10, constructed by Guilford (23) from the results obtained with his standardization group.

In order to extract the actual items which differentiate the groups, an analysis was made on each item in the inventory. The frequency with which the subjects in each group answered each item "yes," "?" or "no" was computed. Since a "?" response indicated an inability to decide between the "yes" and the "no" answer, the frequency of the "?" responses was divided equally between the "yes"

TABLE 2

RELIGIOUS DISTRIBUTION OF THE THREE GROUPS AND GENERAL HOSPITAL POPULATION IN TERMS OF PERCENTAGE IN EACH OF THE THREE MAJOR RELIGIONS

GROUP	N	% CATHOLIC	% JEWISH	% PROTESTANT	TOTAL
Colitis	27	37	56	7	100
Ulcer	30	70	17	13	100
Nonpsychosom.	44	52	32	16	100
Hospital	1424	55	15	28	98

TABLE 3

DIFFERENCES IN IQ SCORES ON THE WECHSLER-BELLEVUE INTELLIGENCE TEST BETWEEN THE THREE GROUPS AND BETWEEN THE NONPSYCHOSOMATIC GROUP AND THE COMBINED PSYCHOSOMATIC (ULCER PLUS COLITIS) GROUP

GROUP	*N*	IQ	*diff*	*t*
Nonpsychosom.	44	111.41	2.74	1.10
Ulcer	30	108.67		
Ulcer	30	108.67	11.55	3.93*
Colitis	27	120.22		
Colitis	27	120.22	8.81	3.36*
Nonpsychosom.	44	111.41		
Nonpsychosom.	44	111.41	2.73	1.18
Combined Psychosom.	57	114.14		

* Significant at .01 level of confidence.

and the "no" responses, thus reducing the responses to two categories. Items which differentiated with critical ratios of 2 or more were then selected.[3]

There were 51 (out of 511 items) which differentiated the duodenal ulcer from the nonpsychosomatic group with critical ratios of 2 or more. There were 56 such items differentiating the ulcerative colitis from the psychosomatic group, and 51 items differentiating the entire psychosomatic group from the nonpsychosomatic group. In all three cases the number of items differentiating the groups is greater than could be expected by chance at the .01 level. There were 34 items differentiating the duodenal ulcer from the ulcerative colitis group which were no more than could be expected by chance. Space limitations do not permit the presentation of the actual items which differentiated the various groups from each other, but they are available to interested readers.[4]

In the Thurstone Interest Schedule the only area of vocational interests which differentiated the three groups was that of biological sciences, in which the colitis patients scored significantly higher than the other two groups. However, when the difference in intelligence between these groups is removed by covariance analysis, the difference in score in biological sciences is reduced to a nonsignificant level. Thus no significant differences in vocational interests were found.

The differences in height, weight, and height/weight ratio between the three groups were not significant.

On the multiple-choice questionnaire, 9 of the 66 items yielded significant results. These items indicate that the ulcer patients reported that they had "strict" or "domineering" fathers, ran away from home one or more times, "disliked" school, "liked" their previous job, tended to go out with people older than themselves, and felt that they had "many friends." The colitis patients tended to have "strict" and "domineering" fathers also, parents who "did not get along at all well with one another" or only "got along moderately well," never ran away from home, "liked" school, and now have few or no "friends."

DISCUSSION

The results on the Wechsler-Bellevue test are in general agreement with the few studies in the literature which deal with the intelligence of these two psychosomatic groups. In a study of ulcerative colitis patients (24), for example, it was found that "80% had an IQ of 115 or more." In the present study 70% of the colitis patients had an IQ of 115 or more. Many writers (3, 6, 14) refer to their colitis patients as being of high intelligence. Mahoney (30), on the basis of Rorschach interpretation, classifies the intelligence of his 20 ulcerative colitis patients as ranging from average to very superior with the median level being above average. Ruesch (38), in a recent study of civilian and Naval personnel with duodenal ulcers, found Wechsler-Bellevue (short form) IQ's of 107 and 111 in his two groups, which compare quite closely with the IQ of the ulcer group, 109, in this study. Osborne and Sanders (36) found a mean IQ of 93 in 15 duodenal ulcer patients by means of the Wechsler-Bellevue scale.

Although there was a significantly larger percentage of Jewish patients in the colitis group than in the other groups or in the general hospital population, no indications ap-

[3] The critical ratios were obtained by means of a graphic method devised by Dr. Joseph Zubin (45) which is based on the ratio of the difference in percentages to the standard error of the difference.

[4] The complete item analysis of the Guilford-Martin inventory as well as the differentiating multiple-choice questionnaire items are on file at the American Documentation Institute. Order Document No. 3840 from ADI Auxiliary Publications, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting $1.00 for microfilm (images 1 inch high on standard 35-mm. motion picture film) or $2.10 for photocopies readable without optical aid.

TABLE 4

RAW SCORES, DIFFERENCES IN SCORES, *t*-TEST RESULTS, AND VARIANCE RATIOS BETWEEN EACH OF THE GROUPS FOR EACH OF 13 FACTORS IN GUILFORD-MARTIN INVENTORY

	N	S	T	D	C	R	G	A	M	I	N	O	Ag	Co
Ulcer	30	16.13	27.50	17.03	20.97	44.77	10.83	20.73	23.10	35.40	24.27	54.77	39.60	60.97
Nonpsych.	44	11.61	27.70	9.25	13.75	44.95	10.80	24.23	24.75	38.34	32.05	57.95	36.23	65.05
diff.		4.52	.20	7.78	7.22	.18	.03	3.50	1.65	2.94	7.78	3.18	3.37	4.08
t		3.13*	.09	4.32*	3.17*	.09	.03	2.58**	1.47	1.87	3.78*	1.16	1.30	1.06
F		1.09	1.51	2.16†	1.68	1.05	1.53	1.18	2.73†	1.51	1.59	1.39	1.38	1.39
Colitis	27	18.33	29.56	16.37	21.70	39.56	9.59	20.74	23.22	33.44	25.74	54.11	35.33	62.78
Nonpsych.	44	11.61	27.70	9.25	13.75	44.95	10.80	24.23	24.75	38.34	32.05	57.95	36.23	65.05
diff.		6.72	1.86	7.12	7.95	5.39	1.21	3.49	1.53	4.90	6.31	3.84	.90	2.27
t		3.95*	.83	3.53*	3.37*	2.33**	1.12	2.31**	1.75	2.88*	3.18*	1.39	.32	.62
t_c¶		3.13		3.41	2.63	1.34		2.02		2.65	2.87			
F		1.63	1.56	2.81†	1.59‡	1.70§	1.72	1.63	1.08	1.74	1.20	1.25	1.05	1.08
Ulcer	30	16.13	27.50	17.03	20.97	44.77	10.83	20.73	23.10	35.40	24.27	54.77	39.60	60.97
Colitis	27	18.33	29.56	16.37	21.70	39.56	9.59	20.74	23.22	33.44	25.74	54.11	35.33	62.78
diff.		2.20	2.06	.66	.73	5.21	1.24	.01	.12	1.96	1.47	.66	4.27	1.81
t		1.23	.75	.25	.23	2.01**	1.11	.01	.09	.95	.60	.20	1.53	.42
F		1.65	1.03	1.30	1.04	1.66	2.65†	1.40	2.79†	1.21	1.32	1.11	1.32	1.51
Psych.	57	17.18	28.47	16.72	21.32	42.30	10.25	20.74	23.16	34.47	24.96	54.46	37.58	61.82
Nonpsych.	44	11.61	27.70	9.25	13.75	44.95	10.80	24.23	24.75	38.34	32.05	57.95	36.23	65.05
diff.		5.57	.77	7.47	7.57	2.65	.55	3.49	1.59	3.87	7.09	3.49	1.35	3.23
t		4.09*	.40	4.43*	3.69*	1.41	.67	2.86*	1.46	2.73*	3.10*	1.50	.61	1.03
F		1.44	1.53	2.43†	1.63	1.45	1.16	1.38	2.61†	1.63	1.40	1.30	1.18	1.05

* *t* significant at .01 level of confidence.
** *t* significant at .05 level of confidence.
† Variance ratio (*F*) significant at either .01 or .05 level of confidence.
‡ Covariance adjustment for differences in intelligence reduces significance to .05 level.
§ Covariance adjustment for differences in intelligence reduces difference in score on factor to a nonsignificant level.
¶ t_c=results of *t* test adjusted for differences in intelligence by covariance on those factors where applicable.

pear in the literature as to the differential incidence of colitis in various racial or religious groups.

There was no significant difference in the occupation level of the three groups which could be ascertained from the classification of occupations into the broad categories listed in the *Dictionary of Occupational Titles* (46). Despite their higher intelligence and somewhat higher educational achievement, the reported income of the colitis patients was approximately the same as the other two groups.

Four factors of the Guilford-Martin inventory yielded significant results at the .01 level of confidence as between the ulcer and the nonpsychosomatic groups. These are S (Social Introversion-Extroversion), D (Depression), C (Cycloid Disposition), and N (Nervousness). One factor, A (Ascendance-Submission), differentiates at the .05 level. On the D factor, the *F* test of the variance ratio between the two groups was significant. Thus the *t* test assumption of equal population variance does not hold between these two groups for this factor, and it cannot be concluded that the differences in D scores are true differences in means. The ulcer group contributes the greater variance to the variance ratio in the M (Masculinity-Femininity) factor as well as in the D factor.

In terms of Guilford's definitions of his factors (23), the ulcer group reported themselves to be more "shy," "withdrawn," "seclusive," tending to have more "cycloid tendencies as shown in strong emotional reactions, fluctuations in mood, disposition toward flightiness and instability"; to be more "socially passive," "jumpy, jittery, with a tendency to be easily distracted, irritated, and annoyed."

Between the ulcer group and the nonpsychosomatic group the number of significant differentiating factors and individual items was greater than could be expected by chance. Since 511 individual items were compared, one would expect approximately 25 items to give significant results ($p \leqq .05$) by chance. Of the 13 factors, by chance one might expect .6 factor to give significant results. Between the ulcer and the nonpsychosomatic groups 4 factors and 51 items differentiated the two groups significantly. By the chi-square test both the number of factors and the number of items which differentiate the two groups are greater than could be expected by chance, at the .01 level. Hence, personality differences, as defined in this study, exist between the ulcer and the nonpsychosomatic groups.

Seven factors yielded significant differences between the colitis and the nonpsychosomatic groups. These are S (Social Introversion-Extroversion), D (Depression), C (Cycloid Disposition), R (Rhathymia), A (Ascendance-Submission), I (Inferiority Feelings), and N (Nervousness). The *F* ratio on the D factor was significant, the colitis group having the greater variance. The differences in scores on the R factor were the only ones to drop to a nonsignificant level when adjusted for differences in intelligence by covariance analysis (Table 4).

In addition to the same factors which differentiated the ulcer patients, in terms of Guilford's definitions, the colitis patients also show a greater tendency to "lack of confidence, underevaluation of one's self, and feelings of inadequacy and inferiority" in comparison with the nonpsychosomatic groups. Again both the number of factors, 5, and the number of individual items, 56, which differentiate the two groups are sufficient to enable us to conclude that there are personality differences between the colitis and the nonpsychosomatic groups ($p = .01$).

When the two groups of ulcer and colitis patients are treated as one group of psychosomatic patients and then compared with the nonpsychosomatic group, significant results occur in 6 factors, in one of which, D, the variance ratio is again significant (Table 4). In comparison with the nonpsychosomatic group, then, the psychosomatic patients, irrespective of diagnosis, report themselves as being, according to Guilford's definitions (23), more "shy," "withdrawn," "seclusive"; tending to have greater "cycloid tendencies as shown in strong emotional reactions, fluctuations in mood, disposition toward flightiness and instability"; more "socially passive"; tending to "lack of confidence, underevaluation of one's self, and feelings of inadequacy and inferiority"; "jumpy, jittery, with a tendency to be easily distracted, irritated and annoyed." On the basis of the number of factors, 5, and the number of individual items, 51, which differentiate these groups, it can be concluded that there are personality

differences between the nonpsychosomatic group and the combined psychosomatic group ($p=.01$).

The only factor which yielded a significant result between the ulcer and the colitis groups was R (Rhathymia), at the .05 level. The colitis patients tend to show a "more inhibited disposition, and an over-control of the emotions." The number of factors, 1, and the number of individual items, 34, which differentiate the two groups are not significantly greater than one would expect by chance. Therefore, personality differences, as defined in this study, were not found between the ulcer and the colitis groups.

A further comparison between the groups of patients can be made in terms of the norms of the Guilford standardization group. When the raw scores are converted into the standard scores based on the standardization group (23), the standard scores of the nonpsychosomatic group are higher than those of either of the psychsomatic groups in 12 of the 13 factors, all but Ag (Agreeableness). Guilford (23) considers that in all but 3 of the factors the higher standard score is indicative of "better adjustment." In the other 3 factors, T, R, and G, he considers a score in the middle range to be most "desirable." Thus in 9 of the 13 factors the tendency of both of the psychosomatic groups is in the direction of the less "well adjusted personality," according to Guilford's criteria. Only on the Ag factor is the score of the ulcer group in the more desirable direction than that of the nonpsychosomatic group.

The personality characteristics of the two psychosomatic groups which were found to differentiate them from the nonpsychosomatic group can be compared with some of the personality characteristics attributed to them by the literature. Although a careful examination of the literature reveals many of the very same adjectives to be descriptive of both groups, there are some descriptions which continually recur in various writers. The tendency to cycloid disposition and fluctuations in mood in ulcer patients has been observed previously (10). Poser (37) found by means of the Rorschach that his patients tended to be more impulsive than his controls. In our study the ulcer group was significantly more "impulsive" than the colitis group but not more so than the nonpsychosomatic group. "Tenseness," "nervousness," and "irritability" have also been attributed to ulcer patients (26, 31). Most writers (1, 4, 38) use such terms as "go-getter," "energetic," "striving," "strongly ambitious," "hard-driving," "self-assertive" in their description of ulcer patients. In the present study the ulcer patients were inclined to passivity. This apparent difference in findings may be reconciled in two ways. First of all, the "strong-ambition" type of personality was noted mainly in "hard-driving" business men. Kapp (26) noted that, although his regular ulcer patients were quite assertive, a group of patients whom he examined on a charity ward were very passive. The subjects in this study would be more akin to the latter group. In their analysis of figure drawings of duodenal ulcer patients, Modell and Potter (32) found them to be basically passive despite their attempt to put up a facade of being more assertive and adequate. Ruesch (38) describes his patients as being "dependent, conforming, overtly counteractive and covertly passive." Second, most of these writers interpret this outward assertiveness as a reaction to a conflict over basic "passive-dependency feelings" (4).

Among the descriptive adjectives applied to the ulcerative colitis patients which this study tended to confirm are "passive," "insecure," "shy," "sensitive," and "moody" (6, 7, 14, 39, 43). On the Rorschach test Mahoney (30) found his group of ulcerative colitis patients to be passive and unable to assert themselves. Brown, Preu, and Sullivan (3) found them to be "lacking in energy." Although their G (general activity) score, which is closest to this attribute, is not significantly different from that of either of the other two groups, it is lower than that of the ulcer and the nonpsychosomatic group.

Since almost the same adjectives are often used to describe both groups, the failure to find significant differences between the ulcer and the colitis groups is not inconsistent with the literature. The general trend of the results in this study seems to point to the existence of strongly similar elements in the two psychosomatic groups, raising the possibility of the existence of a distinguishable "psychosomatic" personality per se, or perhaps a

"gastrointestinal" personality, rather than a specific type for each particular illness. The differences in intelligence between the ulcer and the colitis groups are considered to be important and, if confirmed by other studies, would raise interesting theoretical questions. The individual items which differentiate are at present merely suggestive and need cross validation on other samples of the same populations before they can be offered as clinical tools to aid in the detection of potential ulcer or colitis patients. Such validation studies are now in process.

The need for further studies with psychological tools to establish differences between various psychosomatic groups, especially differences at deeper and more subtle levels of personality, is evident. It is suggested that such studies make careful selection of control groups and more extensive use of groups of patients with demonstrable organic illness of known etiology, but with syndromes of physical symptoms similar to suspected psychosomatic groups, e.g., ulcerative colitis and amoebic dysentery.

SUMMARY

Three groups of patients in a general medical hospital were given a group of objective personality tests to ascertain personality differences.

Two of the groups consisted of patients diagnosed as having diseases generally considered to be in the psychosomatic category, duodenal ulcer and ulcerative colitis. The third group was composed of patients with no history of any psychosomatic disturbances.

1. The ulcerative colitis group attained significantly higher scores on the intelligence test than did either of the other two groups.

2. Personality differences were defined in terms of significant differences in scores on the Guilford-Martin Factor Inventory.

3. Statistically significant personality differences were found between each of the psychosomatic groups and the nonpsychosomatic group, as well as between the combined psychosomatic group and the nonpsychosomatic group.

4. The differences between the ulcer and the colitis groups were not large enough to provide evidence that there were personality differences between these two groups.

REFERENCES

1. ALEXANDER, F. The influence of psychologic factors upon gastro-intestinal disturbances: a symposium. *Psychoanal. Quart.*, 1934, 3, 501–537.
2. ALEXANDER, F., & FRENCH, T. M. *Studies in psychosomatic medicine.* New York: Ronald Press, 1948.
3. BROWN, W., PREU, P., & SULLIVAN, A. J. Ulcerative colitis and the personality. *Amer. J. Psychiat.*, 1938, 95, 407–420.
4. BRUSH, A. L. Recent literature relative to the psychiatric aspects of gastro-intestinal disorders. *Psychosom. Med.*, 1939, 1, 423–428.
5. CATTELL, R. B. *Description and measurement of personality.* New York: World Book Co., 1946.
6. DANIELS, G. E. Psychiatric aspects of ulcerative colitis. *New Eng. J. Med.*, 1942, 226, 178–184.
7. DANIELS, G. E. Nonspecific ulcerative colitis as a psychosomatic disease. *Med. Clin. N. America*, 1944, 28, 593–601.
8. DRAPER, G., DUPERTUIS, C., & COUGHEY, J. *Human constitution in clinical medicine.* New York: Hoeber, 1944.
9. DUNBAR, F. *Emotions and bodily changes.* (3rd Ed.) New York: Columbia Univer. Press, 1946.
10. DUNBAR, F. *Psychosomatic diagnosis.* New York: Hoeber, 1948.
11. EYSENCK, H. J. *Dimensions of personality.* London: Kegan Paul, Trench, Trubner, 1947.
12. EYSENCK, H. J. Review of Guilford-Martin inventory. In O. Buros (Ed.), *The third mental measurements handbook.* New Brunswick: Rutgers Univer. Press, 1949.
13. GARRETT, H. E. *Statistics in psychology and education.* (3rd Ed.) New York: Longmans, Green, 1948.
14. GROEN, J. Psychogenesis and psychotherapy of ulcerative colitis. *Psychosom. Med.*, 1947, 9, 151–174.
15. GUILFORD, J. P. *An inventory of factors STDCR: Manual of directions and norms.* (Rev. Ed.) Beverly Hills, Calif.: Sheridan Supply Co., 1944.
16. GUILFORD, J. P., & GUILFORD, R. B. An analysis of the factors in a typical test of introversion-extroversion. *J. abnorm. soc. Psychol.*, 1934, 28, 377–399.
17. GUILFORD, J. P., & GUILFORD, R. B. Personality factors S, E, and M, and their measurement. *J. Psychol.*, 1936, 2, 109–128.
18. GUILFORD, J. P., & GUILFORD, R. B. Personality factors D, R, T, and A. *J. abnorm. soc. Psychol.*, 1939, 34, 21–36.
19. GUILFORD, J. P., & GUILFORD, R. B. Personality factors N and GD. *J. abnorm. soc. Psychol.*, 1939, 34, 239–248.
20. GUILFORD, J. P., & MARTIN, H. *The Guilford-Martin inventory of factors GAMIN: Manual of directions and norms.* (Abridged Ed.) Beverly Hills, Calif.: Sheridan Supply Co., 1943.
21. GUILFORD, J. P., & MARTIN, H. *The Guilford-Martin personnel inventory: Manual of directions and norms.* Beverly Hills, Calif.: Sheridan Supply Co., 1943.
22. GUILFORD, J. P., & MARTIN, H. Age differences and sex differences in some introvertive and emotional traits. *J. gen. Psychol.*, 1944, 31, 219–229.
23. GUILFORD, J. P., & MARTIN, H. *The Guilford-Martin temperament profile chart.* Beverly Hills, Calif.: Sheridan Supply Co., 1943.
24. HEATH, F. K. (Ed.) Ulcerative colitis—combined staff clinics. *Amer. J. Med.*, 1949, 6, 481–494.
25. KALDEGG, A., & O'NEILL, D. Rorschach patterns in duodenal ulcer. *J. ment. Sci.*, 1951, 96, 190–198.

26. KAPP, F. T., ROSENBAUM, M., & ROMANO, J. Psychological factors in men with peptic ulcer. *Amer. J. Psychiat.*, 1946, 103, 700–704.
27. KORNREICH, M. Variations in the consistency of the behavioral meaning of personality test scores. *Genet. Psychol. Monogr.*, 1953, 47, 73–138.
28. LINDQUIST, E. F. *Statistical analysis in education research.* Boston: Houghton Mifflin, 1940.
29. LOVELL, C. A study of the factor structure of thirteen personality variables. *Educ. psychol. Measmt.*, 1945, 5, 335–350.
30. MAHONEY, V. P., BOCKUS, H. L., INGRAM, M., HUNDLEY, J. W., & YASKIN, S. C. Studies in ulcerative colitis. *Gastroenterol.*, 1949, 13, 547–563.
31. MITTLEMANN, B., & WOLFF, H. A. Emotions and gastro-duodenal function. *Psychosom. Med.*, 1942, 4, 5–61.
32. MODELL, A. H., & POTTER, H. W. Human figure drawings of patients with essential hypertension, peptic ulcer, and bronchial asthma. *Psychosom. Med.*, 1949, 11, 282–292.
33. MOSCHOWITZ, E., & ROUDIN, H. The association of psychosomatic disorders and their relation to personality types in the same individuals. *N. Y. State J. Med.*, 1948, 48, 1375–1385.
34. MOSIER, C. I. A factor analysis of certain neurotic symptoms. *Psychometrika*, 1937, 2, 263–286.
35. NORTH, R. D. An analysis of the personality dimensions of introversion-extroversion. *J. Pers.*, 1949, 17, 352–367.
36. OSBORNE, R. T., & SANDERS, W. A. Rorschach characteristics of duodenal ulcer patients. *J. clin. Psychol.*, 1950, 6, 258–262.
37. POSER, E. C. Personality factors in patients with duodenal ulcer: A Rorschach study. *J. proj. Tech.*, 1951, 15, 131–143.
38. RUESCH, J. *Duodenal ulcer.* Los Angeles: Univer. California Press, 1948.
39. SULLIVAN, A. J. Psychogenic factors in ulcerative colitis. *Amer. J. digest. Dis.*, 1936, 2, 651–656.
40. THURSTONE, L. L. *Thurstone interest schedule.* New York: Psychological Corp., 1948.
41. WEISS, E., & ENGLISH, O. S. *Psychosomatic medicine.* (2nd Ed.) Philadelphia: Saunders, 1949.
42. WECHSLER, D. The measurement of adult intelligence. (3rd Ed.) Baltimore: Williams and Wilkins, 1944.
43. WITTKOWER, E. Ulcerative colitis: Personality studies. *Brit. med. J.*, 1938, 2, 1356–1360.
44. WOLFLE, D. Factor analysis in the study of personality. *J. abnorm. soc. Psychol.*, 1942, 37, 393–397.
45. ZUBIN, J. Nomographs for determining the significance of the difference between the frequency of events in two contrasted series or groups. *J. Amer. stat. Ass.*, 1939, 34, 539–544.
46. *Dictionary of occupational titles.* Part IV. Entry occupational classification. Washington, D. C.: War Manpower Comm., 1944.

Received March 24, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

EFFECTS OF INITIAL SERIAL POSITION AND ATTITUDE UPON RECALL UNDER CONDITIONS OF LOW MOTIVATION

LEONARD W. DOOB
Yale University

Most investigators studying the effect of serial position and attitude upon recall require *Ss* to learn certain materials, and motivate them to do so by indicating (2) or eventually demonstrating (4) that they will be subsequently tested or by announcing in advance (3) that some subsequent activity based upon the materials will be demanded. When people are exposed to a mass medium of communication in a naturalistic setting, however, they usually are not motivated directly either to perceive or to learn its contents. Under such conditions, this study asks, how is recall affected by initial serial position? Are items that are presented first (primacy) or last (recency) recalled more effectively? Does an intense or a favorable attitude promote recall?

If a drive-reduction hypothesis be assumed, two propositions may be formulated. In the first place, recall may be considered a function of drive strength at the time of the original perception: *the stronger the drives that are evoked* by prior instructions, by some prior set, or by the stimuli themselves, *the greater the recall.* Then, second, recall also depends upon retroactive inhibition: (*a*) those stimuli that are presented last in a series (recency) are subject to less inhibition than those coming earlier in the series; (*b*) intervening activity is likely to be less disruptive of a more than of a less strongly reinforced response, and the degree of reinforcement is affected by the strength of the evoked drives. Since retroactive inhibition is quantitatively less for those stimuli coming last and since it is more disruptive when the evoked drives are weak, it follows that *the factor of recency more effectively promotes the recall of stimuli that evoke weak rather than strong drives.* The present investigation seeks to explore the implications of these two propositions.

PROCEDURE

Male students in a psychology class were exposed to two different types of situations which differed from each other with respect to motivation-to-perceive but which offered no direct motive-to-learn. The first consisted of a questionnaire containing five prose items. Each item was composed of three sentences totalling about 70 words. The paragraphs offered opinions on five controversial subjects: the function of a liberal arts education, the Tennessee Valley Authority, advertising, religion, and war between the United States and the Soviet Union. The last one, for example, read:

> A war between the United States and the Soviet Union is completely inevitable. High Russian officials are convinced that their country will be secure and that it can progress only when every nation in the world possesses a communist form of government. The United States is the one great power which is opposing, can continue to oppose, and should oppose effectively this imperialist, ruthless, revolutionary ambition.

The class was randomly divided into five groups, each of which received a different item in the first position; the remaining four paragraphs then followed in the same sequence for all of the groups. In order to increase motivation to perceive but not to learn the items and also to evoke and obtain a measure of attitude, these instructions were given: "Read through the following statements and after each indicate with a cross your degree of agreement or disagreement with the point of view." A five-point scale followed each statement (fully agree, agree, disagree, fully disagree, undecided).

Two days later the 357 *Ss* who were in class on both days were asked to perform two tasks:

> 1. The day before yesterday you read through some statements on various topics. In the space below, please LIST the topics you remember at this time.
> 2. For each topic you have listed above, please recall as many of the arguments and ideas as you can which were advanced in the original statements.

The writer and an assistant independently coded the data. They agreed in 96 per cent of the instances for the first task; the remaining 4 per cent were discussed until agreement was obtained. A content analysis was made of the "arguments and ideas" which were recalled from the three sentences of each of the five items: agreement dropped to 83 per cent, but again informal and amicable discussion resolved the differences.

The second situation was a naturalistic one in which no instruction to perceive or learn was provided. Three months after the questionnaire, the same class was asked to see a featured picture at one of the local theaters because, it was said, the picture's communication techniques would subsequently be discussed by the instructor. Before the

lecture on a Thursday, the students were given these instructions:

> We agreed to see *That Wonderful Urge*. If you went, please answer these questions:
>
> 1. When did you go to see the picture: on Saturday, Sunday, Monday, Tuesday, or Wednesday?
> 2. When you sat down in the theater: (a) which picture was being shown? (b) was it at the beginning, the middle, or the end of that picture?
> 3. Finally, if you saw the newsreel, please list the items which that newsreel contained.

The total number of *S*s was reduced to 137 because some did not attend the performance, others had seen the feature picture elsewhere, and many sagacious geniuses timed their arrival at the theater so skillfully that they avoided the atrocious co-feature but also, unfortunately, the newsreel. Seven separate items were shown in the brief newsreel. The writer and an assistant independently coded the data that were recalled, agreed in 93 per cent of the cases, and then reconciled their differences.

RESULTS

The principal results from both situations have been grouped into Tables 1 and 2 which reveal only the statistically significant differences. Breakdowns of a column's variable in terms of a variable from another column are designated under an italicized subheading which is the variable from that other column. Combinations of subgroups are indicated by parentheses and plus signs. The results will be presented in terms of the two propositions which have been suggested in the introduction.

Relation between Initial Drive Strength and Recall

The strength of the drives that were evoked in one or both situations may be viewed in three different ways:

The drive strength of the initial perception. Before perceiving the five prose items, *S*s were motivated not to learn but at least to judge their contents. In the motion-picture situation, on the other hand, they were motivated neither to perceive nor to remember the newsreel since they had been instructed only to see the featured picture. Aside from other differences between the two situations, it would appear that the drives to perceive the prose were stronger than those to perceive the newsreel; hence recall should have been greater in the first than in the second situation. Analysis of the results reveals that only 2.5 per cent of the *S*s failed to recall a single one of the prose items and 7 per cent recalled all five; whereas 29 per cent failed to recall a single newsreel item, 0.7 per cent recalled five, and nobody recalled all seven ($p < .01$ for the difference between the two means, 2.89 and 1.03). Memorizing ability was probably not involved since the r between number of prose and newsreel items recalled was low (.11) and not significant. The same difference appears when the interval between initial exposure and the test situation is equated by comparing the mean number of prose items (2.89) recalled by the total group and the mean number of newsreel items (1.59) recalled by those who had perceived the pictures two days before the recall ($p < .01$), or even the mean number (2.33) recalled by those who had seen it the night before ($p < .05$). These differences are striking since no correction has been introduced to take into account the fact that a total of seven newsreel items but only five prose items could be recalled. Each of the five prose items, moreover, was recalled significantly better than any one of the seven newsreel items ($p < .01$).

The drive strength of the evoked attitudes. The subjects were asked to rate the intensity of their attitudes toward the five prose items on what is essentially a two-point scale: unless they were "undecided," they checked either (*a*) fully agree (or fully disagree) or (*b*) agree (or disagree). The significant differences are listed in column two of Table 1. A brief description of the three measures follows:

1. Percentage of recalled instances. Since five items were presented to 357 *S*s, the total number of instances is 1,785. Of these, 57 per cent were recalled. Ten per cent of the instances originally checked were fully agree, 36 per cent agree, 36 per cent disagree, 11 per cent fully disagree, and 7 per cent undecided attitudes.

2. Mean order of recall. Means are calculated by assigning 1 to items recalled first, 2 to those recalled second, etc.

3. Accuracy of recall. The unit is the sentence, of which there were 15 (three for each of the five items). Table 1 employs an abbreviated but sufficiently precise version of the content analysis that was applied to each sentence of those items whose scope had been correctly recalled:

a. "Correct": correct recall of the sentence's sense but not necessarily its exact wording.

TABLE 1

SIGNIFICANT DIFFERENCES IN RECALL OF FIVE PROSE ITEMS

MEASURE	ATTITUDE INTENSITY	ATTITUDE CONTENT	SERIAL POSITION	
Percentage of recalled instances	† f. agree > agree	undec. > agree	• 1 > 3	• 1 > 4
	position 4	• undec. > (agree + disagree)	• 1 > 5	2 > 3
	(f. agree + f. disagree) > (agree + disagree)	*position 3* (disagree + f. disagree) > (f. agree + agree)	*f. agree + agree* • 1 > 3	*disagree + f. disagree* • 1 > 4
	positions (2 + 3 + 4 + 5 (f. agree + f. disagree) > (agree + disagree)		2 > 3	• 1 > 5
			1 > (2 + 3 + 4 + 5)	1 > 3
			f. agree + f. disagree	1 > (2 + 3 + 4 + 5)
			• 1 > (2 + 3 + 4 + 5)	*undecided*
				1 > (2 + 3 + 4 + 5)
			agree + disagree	
			• 1 > 3	• 1 > 4
			• 1 > 5	1 > (2 + 3 + 4 + 5)
Mean order of recall	none	none	• 1 > 2	• 1 > 3
			• 1 > 4	• 1 > 5
Recall accuracy				
a. % correct	none	(f. agree + agree) > (disagree + f. disagree)	none	
b. % incorrect	none	• (disagree + f. disagree) > undecided	none	
c. % separately omitted	none	none	none	
d. % jointly omitted	none	undecided > (disagree + f. disagree)	none	

• Significant at .01 level; remainder significant at .05 level.
† f.=fully.

b. "Incorrect": a change or omission of the sentence's qualification, a minor or major distortion of its meaning, or a combination of the two.

c. "Separately omitted": no reference to the sentence although parts of one or two of the other two sentences from the same item were recalled.

d. "Jointly omitted": no recall of any of the three sentences from the item, although its scope had been correctly listed in answering the previous question.

In the second column of Table 1 it can be observed that greater recall seems to be related to greater intensity of attitude. When all the instances are considered together, the following percentages of recalled instances are associated with the indicated attitudes: fully agree, 62 per cent; agree, 53 per cent; disagree, 57 per cent; fully disagree, 60 per cent; and undecided, 63 per cent. As shown in that table, however, only the difference between the "fully agree" and the "agree," and not that between the "disagree" and the "fully disagree" instances are statistically significant. Under the serial position specified

in the same column of the table, the combination of "fully agree" and "fully disagree" instances leads to reliably more recall than the combination of "agree" and "disagree." When each of the five prose items rather than instances is examined, it is found (*a*) that for three of them the intensity combination of "fully agree" and "fully disagree" is associated with more—but not significantly so—recall than that of "agree" and "disagree"; and (*b*) that on the fourth item the "disagree" attitude and on the fifth the "fully agree" attitude are significantly associated, respectively, with fewer and more recalls than the other attitudes ($p < .05$). Table 1 also shows that attitude intensity has no relation to the mean order of recall or the accuracy of recall; nor is there an association with the mean number of words that were used in recalling the items (the mean of all instances is 16.2 words).

The content of the evoked attitude. It might be assumed, in the first place, that the *S*s reacted more violently to those five items with which they disagreed than to those with which they agreed; hence the initial drive strength in the former situation would be greater than in the latter, which would mean that more "disagree" items ought to be recalled than "agree" items. It might also be assumed—and this is the so-called Freudian assumption—that the "agree" items during the test situation, though perhaps weaker, would facilitate learning since "congruent" items of this type have led in the past to drive reduction; hence more "agree" items ought to be recalled than "disagree" items. If both assumptions are correct, then the two tendencies might counteract and perhaps cancel each other. The third column of Table 1 indicates no significant differences in recall with respect to agreement and disagreement (except for position 3 which goes counter to the Freudian assumption) so far as "percentage of recalled instances" and "mean order of recall" are concerned. There are also no significant differences with respect to mean number of words used in recall.

For the interval between initial exposure and the test situation, however, the two assumptions might lead to the same anticipation. "Disagree" items may have initially provoked stronger drives according to the first assumption, but the content of what was then learned could be relatively quickly forgotten since no motivation to recall was provided either by those items or by the experimenter. The Freudian assumption suggests that the content of "agree" items was likely to be retained since the items themselves provided some incentive to do so as a result of past drive reduction. The "% correct" row of Table 1, column 3, indicates that "agree" items, when they are recalled, tend to be recalled more accurately than "disagree" items.

Both assumptions, finally, might also lead to a similar anticipation for "undecided" items. The *S*s who were "undecided" may have been thrown into a state of conflict and may have taken longer to reach no decision; the drive strength in the exposure situation was great and hence the items should be recalled. In the test situation, however, no incentive to recall their precise contents may have been present since indecision in general is unrewarding. The third column of Table 1 supplies some confirmatory evidence: "undecided" items are recalled with relatively great frequency, and their contents tend to be recalled neither "correctly" nor "incorrectly" but to be "jointly omitted."

Relation of Serial Position and Recall

It has been suggested that the drives evoked in the prose situation were probably stronger than those evoked in the newsreel situation since perception in the former situation was motivated by the instructions and in the latter no such external motivation was present. Empirically, the recall of the prose items has been shown to be greater than that of the newsreel items. Accordingly, it may be anticipated that those newsreel items whose serial position was last (and hence more "recent") would be recalled more effectively than those coming first in the sequence, but that this factor of recency would not operate for the prose items. In general, the results support the anticipation.

When the percentages of *S*s recalling each newsreel item are arranged in order of their initial serial position, this sequence emerges: (1) 6, (2) 4, (3) 22, (4) 5, (5) 23, (6) 18, and (7) 25 per cent. Items 1, 2, and 4 do not differ significantly from one another, nor do items 3, 5, 6, and 7; but each one in the first group differs from each in the second beyond the .01 level. Naturally the interest

in and the drives evoked by each item must also have affected recall, and perhaps this factor accounts for the fact that more *Ss* recall the third than the second item. On a priori grounds, however, it is difficult to imagine why the funeral of an Egyptian leader, a beauty contest in New York, a parade in Philadelphia, and airplane students over Miami (items 3, 5, 6, and 7, respectively) should appeal to undergraduates more than an art-student ball in London, some lion cubs in Poland, and a windstorm off the English coast (items 1, 2, and 4, respectively). In addition, there is no significant relation between serial position and either the order of recall or the mean number of words used in recalling an item.

The effects of retroactive inhibition upon the recall of the newsreel items can also be measured with respect to two additional temporal factors: the number of days elapsing between exposure to the newsreel and the test situation; and the position of the newsreel on the program when the *Ss* viewed it. The significant differences are listed in Table 2. In the column headed "Number of Elapsed Days," 1 means that *S* saw the newsreel the day before answering the questionnaire, 2 that he saw it two days before, etc. The total group has been dichotomized for purposes of the last column in Table 2, which refers to the position on the program which the newsreel occupied when *S* reported that he saw it: one group saw it toward the end (i.e., shortly before leaving the theater) and the other toward the start (i.e., before seeing the featured and/or the co-featured picture). Retroactive inhibition in the period between the exposure and the test situation increased, it is assumed, as the number of days between the exposure and the test increased; and retroactive inhibition produced by other pictures on the theater's program was greater for those seeing the newsreel at the start than at the end of the program. The measures in Table 2 are self-explanatory except for the "incorrect" recalls: these refer to the fact that 32 per cent of the 137 *Ss* mistakenly recalled one or more newsreel items which had not appeared on the program in question and which, therefore, they had recently seen as part of another program in a different theater.

TABLE 2
SIGNIFICANT DIFFERENCES IN RECALL OF SEVEN NEWSREEL ITEMS

MEASURE	NUMBER OF ELAPSED DAYS	POSITION OF NEWSREEL ON PROGRAM
Percentage of *Ss* recalling one or more items:		
a. correctly	$^{*}(1+2+3) > (4+5)$	end > start
b. incorrectly	$3 > 1$ $\quad 5 > 1$	end > start
Mean number of items recalled:		
a. correctly	$^{*}(1+2+3) > (4+5)$ *end*	none
	$^{*}(1+2+3) > (4+5)$ *start*	
b. incorrectly	$(1+2+3) > (4+5)$	
	$(4+5) > (1+2+3)$	none
Mean words used in recall	$(1+2+3) > (4+5)$ *end*	none
	$^{*}(1+2+3) > (4+5)$	

* Significant at .01 level; remainder significant at .05 level.

Table 2 indicates that, when the data are grouped, there is a strong association between recency and greater as well as more accurate recall. The shorter the time interval between seeing the film and answering the questionnaire, the greater the tendency to recall items correctly and to use more words to describe them, and the less the tendency to recall incorrect items. This tendency is also true of each of the seven items, but only three of the differences are significant between the .01 and .05 levels. The effects of "position of newsreel on program" are less clear-cut. A higher percentage of one or more correct recalls occurs among those seeing the items toward the end than among those seeing them

toward the start of the performance, but there are no differences with respect to the mean number of items recalled. This tendency is true of five of the seven items when they are analyzed separately, but only the differences between the two groups on the item presented first are significant. Going counter to this trend, however, is a tendency for more incorrect recalls likewise to be associated with end rather than with start perception.

In contrast, it may be observed in the last column of Table 1 that primacy rather than recency seems to be associated with the percentages of prose instances recalled and with the mean order of recall. The percentages of recalled instances are the following which are arranged in their original serial position: (1) 66, (2) 59, (3) 52, (4) 54, and (5) 54 per cent. The mean recall positions of all instances are the following which are again arranged in their original serial position: (1) 1.83, (2) 2.26, (3) 2.27, (4) 2.33, and (5) 2.23. The relation between serial position and recalled instances remains under varying combinations of attitude content and intensity. It also tends to be true when the effect of serial position upon each of the five separate items is considered: for four out of five items, first presentations are recalled in greater numbers than the combinations of the remaining serial positions (three of which are either close to or beyond the .01 level) and in the case of the fifth item the percentage of recall is the same among first and subsequent positions. An item coming first perhaps produced the greatest tension, and hence the greatest drive reduction occurred after the attitude scale was checked: maximum concentration to implement the instruction was initially required, but thereafter the task could become more routinized. This interpretation receives some support from the way in which the three sentences composing each item were recalled: more first or third sentences tended to be separately or jointly omitted ($p > .01$), but more correct and fewer incorrect recalls were associated with first than with second or third sentences ($.05 > p > .01$).

DISCUSSION

The results of this investigation seem in general to indicate that recall under conditions of low motivation is affected by the strength of the drives that are evoked and by retroactive inhibition. In evaluating the obtained differences, however, it is well to remember that their magnitude is seldom very great; and that many of them appear only when the data are grouped and hence tend to disappear or become nonsignificant when individual items are considered. In addition, the period of time that elapsed between initial presentation and recall was relatively brief, one to five days. Eberhart and Bauer (1) have indicated some of the complications which arise when a longer period of time, two years, is involved.

The effect of attitude and serial position upon recall under these conditions of low motivation is not consistent. For the five prose items, attitude intensity and undecided attitudes are associated only with percentage of instances recalled, and attitude content with recall accuracy. Items coming first in the prose situation tend to be recalled in greater numbers and to be recalled sooner, but they are not recalled more accurately or with a greater number of words. Items coming last in the newsreel situation tend to be recalled in greater numbers, but they are not recalled sooner or with a greater number of words. The mean number of words used in recall, however, is associated with the number of days elapsing between exposure and recall of the newsreel items, although not with the position of the reel on the program.

When motivation to learn is low, therefore, no single principle is likely to provide the basis for predicting the quality and quantity of recall. Primacy, for example, may be of some significance but only because of the relatively strong drive that is evoked, and then order of recall may be affected. Recency may be important, but only when drive strength to perceive and to learn is relatively weak, and then order of recall may not be affected. Regardless of the content of their attitudes, people may remember that which they have experienced intensely—or, as Watson and Hartmann (4) have shown, that which they consider "effective"—but their recall may not be very accurate or detailed. People may remember what they wish to remember, but only when a drive provides the incentive for recall in some form to occur. Quite obviously the variables of drive strength and retroactive inhibition interact with each

other and with other variables that must be more clearly identified.

SUMMARY

Three hundred and fifty-seven undergraduates indicated their attitude toward five pieces of controversial prose which they received in five different serial orders. Without prior warning they attempted two days later to recall the scope and the contents of the items. From the same group, 137 witnessed a newsreel in a regular theater program and then, also without prior warning, they were asked from one to five days later to recall the items.

1. Recall of the prose items is clearly superior to that of the newsreel items.

2. In the case of the prose items, quantity of recall tends to be associated with attitude intensity, with undecided attitudes, and with primacy in the original serial position; and quality of recall with congruent attitudes. Recall order is associated only with primacy.

3. In the case of the newsreel items, quantity of recall tends to be associated with recency, but greater recall of items that were not on the program—incorrect recalls—is produced by the group which saw the newsreel later on the program.

4. The results are presented and discussed in terms of drive strength and retroactive inhibition, and it is concluded that more than one psychological principle is required to explain the factors affecting recall under conditions of low motivation.

REFERENCES

1. EBERHART, J. C., & BAUER, R. A. An analysis of the influences on recall of a controversial event: the Chicago *Tribune* and the Republic Steel strike. *J. soc. Psychol.*, 1941, **14**, 211–228.
2. EDWARDS, A. L. Political frames of reference as a factor influencing recognition. *J. abnorm. soc. Psychol.*, 1941, **36**, 34–50.
3. GUEST, L. P., & BROWN, R. H. A study of the recall of radio advertising material. *J. Psychol.*, 1939, **8**, 381–387.
4. LEVINE, J. M., & MURPHY, G. The learning and forgetting of controversial material. *J. abnorm. soc. Psychol.*, 1943, **38**, 507–517.

Received May 1, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

RELATION OF BACKGROUND EXPERIENCE TO SOCIAL ACCEPTANCE [1]

MORTIMER R. FEINBERG
City College of New York, and Richardson, Bellows, Henry & Co.

ACCEPTANCE by the social group is an important need of the individual in our culture. The relationship between this social acceptance and specific life experiences, however, is little understood. This research reports findings comparing sociometric acceptance with the background experiences of adolescent boys at three economic levels.

Psychologists in almost every field of specialization have placed great emphasis on background experiences in the study of personality development. The social worker emphasizes the case history, the clinical psychologist stresses environmental factors in the etiology of neuroses and psychoses, and the personnel psychologist utilizes case history items in the application blank. However, it is unfortunately true that some psychologists will interpret certain background experiences as being particularly important in the creation of some given ill-defined behavior pattern solely on the basis of their adherence to a theoretical system.

Recent progress has been made in both civilian and military research, particularly in the latter, where background data have been treated more objectively and have been related to such criteria as leadership (16), occupational success (3, 4, 8, 12, 15), and personality adjustment (2, 7, 10). The most systematic quantification of background data was accomplished by the Personnel Research Section, AGO (14) in their development of the Biographical Information Blank, which is employed in the selection of officers for retention in the peacetime Army.

However, there have been very few quantitative attempts to correlate background data with the criterion of social acceptance. The success of Dale Carnegie's books, the ability of so many to identify, in *Death of a Salesman*, with Willy Loman's hunger for approval, the hypothesis of Riesman (11) that there has been a shift in America from dominance by the "inner directed" type of person to dominance by an "outer directed" type to whom personal goals are less significant than esteem by others—all these attest to the importance of social approbation in our culture. This need for "popularity" is also strong at the adolescent level, where all the magazines for teenagers attempt to give rules by which the young person can guide himself along the path to social recognition.

The relationship between family background and favorable or unfavorable personality development (1, 5, 9, 10) has been investigated. Probably the most extensive of these studies dealt with socioeconomic differences and their effect on adolescent personality development (6).

We know of no previous attempts, with large samples, to quantify background data at the adolescent level and systematically relate it to social acceptance. Also, we know of few attempts to compare quantified background information among contrasting economic levels, and almost no attempt to investigate the commonness in predictive background experiences among the different economic levels.

The major emphasis in this article will be on an empirical analysis of the items significant in predicting the criterion of social acceptance. The development of the scoring keys and the validity and reliability of the Personal History Questionnaire will be discussed briefly here and described in more detail in a future paper.

[1] This study was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at New York University Graduate School of Arts and Science. The author wishes to express his gratitude to the staff of Richardson, Bellows, Henry, and Company, Inc., particularly Dr. Marion W. Richardson, who have helped greatly as statistical advisers. He desires to acknowledge the assistance of Dr. David Segel of the Federal Security Agency, Department of Education, in the obtaining of the Indiana sample, and Dr. Sheldon S. Zalkind and Dr. Max Smith for valuable editorial suggestions.

A special debt of gratitude is due Professor Douglas H. Fryer, under whose direction this thesis was undertaken and under whose guidance and encouragement it was completed.

METHOD

When analyzing the effect of background experiences on human subjects, it is rarely possible to control our variables, split our groups, and institute the controls necessary for experimental study. Therefore, for the purposes of this research, techniques have been borrowed from the area of tests and measurements. This approach can only indicate concomitant variation and not causality.

Variables

A Personal History Questionnaire[2] was developed, consisting of 240 items, written to represent six areas of adolescent activity. Items concerned with the areas of family and home experiences, intellectual activities, school experiences, social activities, degree of economic independence, and adjustment to concept of self were included.

Criterion

A modified "buddy rating" or sociometric technique was used as the criterion of social acceptance. The pupils nominated the four boys in their class whom they would like to have sitting next to them the following term—who in their opinion would make the most desirable friends; and also the four boys with whom they would feel uncomfortable, who give them a feeling of uneasiness or annoyance in the classroom situation.[3] The class groups in which this was accomplished approximated 30 boys.

Economic Level

Since a large sample was required, an economic variable relatively easy to obtain was selected for comparing samples subjected to influences associated with three economic levels.

The term *economic level* as used here refers to relative financial status of the family. The samples selected according to economic status are described as follows:

High economic level. Schools with boys whose estimated family income was above $9,000 a year.

Middle economic level. Schools with boys whose estimated family income was approximately $5,000 to $7,500 a year.

Low economic level. Schools with boys whose estimated family income was below $3,000 a year.

Great care was taken in drawing the samples from economic level so as to make certain that there would be almost no overlap between levels. The middle and low economic levels were taken from the public schools of New York City by first consulting an official of the Board of Education who assigned an economic level to the different schools on the basis of neighborhood, reports from census figures, etc. Second, the principals of the particular schools classified their own students as belonging in the main to a particular economic level. There was complete agreement between the officials' and the principals' estimates. While there may be a few students in each of the selected schools who do not belong to the economic level of the majority of their classmates, their number is probably insignificant. If any contamination of specific economic levels has occurred, it is more likely to be between the middle and low economic levels since the samples of the high economic level were drawn exclusively from private schools with high tuition rates.

Subjects

The total sample tested numbered over 2,000 adolescent boys between the ages of 13 and 16 years, residing in New York City, in a suburb of New York City (Westchester County), and in a small county in Indiana. In the selection of samples, emphasis was placed upon adequate representation for one area of the country, although it was recognized that there might be differences in predictive background experiences for individuals living in different parts of the country. Six samples of approximately 300 subjects (*Ss*) each were secured in New York City, two at each economic level. The first sample at each level was obtained in May, 1948, and the second in November, 1948. The same schools but different *Ss* were used in securing these samples. The samples outside of New York City consisted of 300 *Ss* from a Westchester County school and 100 *Ss* from an Indiana school.

DATA TREATMENT

Determining the Criterion Score of Social Acceptance

The students' sociometric choices were scored in class groups by means of a scoring chart devised for the purpose. For each student the number of rejections was subtracted from the number of acceptances. The difference, positive or negative, was divided by the total number of individuals in the class and multiplied by a hundred. Based on this acceptance-rejection score, the *Ss* were divided into high, middle, and low criterion groups. If an *S*'s criterion score was 15 or higher, he was placed in the high criterion group; if he received more rejections than acceptances he was placed in the low group; and any *S* with a score falling between these two was placed in the middle criterion group. Those *Ss* who obtained a score of zero (receiving no acceptances or rejections, or an equal number of

[2] The questionnaire was developed with the cooperation of Dr. D. H. Fryer, Dr. E. H. Rundquist, and Mr. E. R. Danziger. The original items and validated questionnaire with manual can be obtained from Richardson, Bellows, Henry, and Company, 1 West 57th Street, New York, New York.

[3] The *Ss* also listed the adjectives which were, in their opinion, descriptive of the behavior of the boys they accepted or rejected. These adjective descriptions have been analyzed for each economic level and the results will be discussed in a future paper.

acceptances and rejections) were eliminated from the analysis. This was done to purify the criterion and it is also felt that such *S*s should be studied as a separate group.

Method of Analyzing Responses to Personal History Questionnaire

Hollerith (IBM) cards were punched for each *S* in the first and second samples of the New York City population. These cards coded all the necessary information—the *S*'s number, class, school, criterion score, and his answers to the questionnaire. Approximately 8,000 cards were punched, four for each *S*. Cards were then sorted for frequency of response to each of the 240 item alternatives of the questionnaire according to each criterion group within each economic level.

Method of Computing Significance of Differences

Having selected the criterion groups and having obtained item frequencies for the criterion groups on each item alternative, the statistical significance of differences between the responses of accepted and rejected boys was determined by the chi-square statistic. The *p* values were computed only when the frequency for each alternative exceeded ten. Also, for each item, the percentage of responses to each alternative in the high, middle, and low criterion groups was computed. The first test used for acceptance of an alternative for comparison of economic levels, or inclusion in a scoring key, was that the response percentage in the middle criterion group should fall between those of the high and low criterion groups. This was done to establish significant alternatives in linear relationships on the criterion of social acceptance. The second test for the acceptance of an alternative was that the chi square between the high and low criterion groups yield a probability (*p*) value at the .20 level. This critical level is often used by experts in test construction.

Method of Constructing Scoring Keys and Determining Validity of Prediction

Keys for scoring the questionnaires of each economic group were developed, permitting a quantified measure of the relationship between background experiences and the criterion. Six separate keys were constructed for this purpose from the New York City population, one for each economic level of the first and second samples. Later, those items which were not similarly predictive of the criterion for both samples at the same economic level were eliminated, establishing one final key for each economic level. An over-all key, applicable to all economic levels, was constructed by retaining only those items which were statistically significant in three out of the six economic samples.

This final over-all key, consisting of 45 items, was applied to the questionnaires secured from a population living in a suburban area of Westchester County, New York, and one from a midwestern farm area of Indiana. From this cross validating, a tentative conclusion can be drawn as to the geographical commonness of predictive background experiences of adolescents and the extent to which it is possible to predict the criterion in one locality with scoring keys devised in another area. No attempt was made in the samples outside of New York City to differentiate economic levels.

All correlations reported between the questionnaire scores and the criterion scores were Pearson coefficients.

A Method of Comparison between Economic Groups by Significant Items

Item significance in predicting the criterion is, of course, included in the scoring procedure indicated above. But it was believed worth while as a part of the problem to indicate items of background and classes of such items that were consistent or were in opposition among economic levels.

RESULTS

Analysis of Questionnaire (Scoring Keys, Reliability, Validity)

Correlations between scores on the Personal History Questionnaire and criterion scores of social acceptance resulting from an application of the final scoring key for the three economic levels are shown in Table 1. Table 2 shows correlation by economic level when the Personal History Questionnaires were scored with the over-all scoring key constructed from all economic samples. Fig. 1 compares the correlation coefficients

TABLE 1

CORRELATION BETWEEN QUESTIONNAIRE SCORES AND SOCIOMETRIC CRITERION AT THREE ECONOMIC LEVELS IN NEW YORK CITY

SAMPLE	LOW ECONOMIC KEY *		MIDDLE ECONOMIC KEY *		HIGH ECONOMIC KEY *	
	CORRELATION	N	CORRELATION	N	CORRELATION	N
First	.66±.40	204	.73±.033	196	.68±.031	308
Second	.67±.037	235	.72±.031	246	.68±.034	246
Combined	.66±.029	439	.72±.024	442	.68±.022	554

* Final key based on first and second sample.

obtained by an application of these two keys. It will be recalled that the final over-all key was constructed from items significant in three out of the six samples. Therefore, an alternative of an item could be included in the over-all key if it was significant for only two economic levels; for example, it might be significant in both middle economic samples and in only one of the high economic samples. In such an instance the influence of the low economic group would not be present in the scoring of the item.

If there were a large overlap of predictive background among the economic levels it would be expected that an over-all key would predict the criterion about as well as one devised specifically for one economical level. This is substantially what occurred. Only at the low economic level (see Fig. 1) was the correlation obtained from the use of the final low economic key as much as .10 points greater than that obtained from the use of the final over-all key. For the middle and high economic levels there was very little difference.

In summary, it can be stated that there is considerable commonness of adolescent predictive background experiences among different economic levels. It appears that, for measurement purposes, one can predict social acceptance of boys by their associates by means of background experiences with almost equal accuracy without a knowledge of their economic level as one can with this knowledge.

The over-all scoring key, which yielded a minimum correlation of .66 for the New York population, was used in scoring the Personal History Questionnaires of two samples outside of New York City. As shown in Table 3, the correlation coefficient obtained for a Westchester sample was .48, which is regarded as highly significant for a predictor of such a criterion as social acceptance. The Indiana sample showed a definite loss in prediction and the correlation (.25) was below the level of significance useful for a single predictor. The small number of boys in this sample and the marked homogeneity of the group which was from a rural population may have been influences determining the lower correlation.

Analysis of Items

Common background experiences predicting the criterion at all economic levels. The

TABLE 2

CORRELATION BETWEEN QUESTIONNAIRE SCORES AND SOCIOMETRIC CRITERION AT THREE ECONOMIC LEVELS IN NEW YORK CITY

SAMPLE	LOW ECONOMIC LEVEL *		MIDDLE ECONOMIC LEVEL *		HIGH ECONOMIC LEVEL *	
	CORRELATION	N	CORRELATION	N	CORRELATION	N
First	.53±.051	204	.79±.029	196	.70±.030	308
Second	.59±.043	235	.62±.040	246	.56±.044	246
Combined	.55±.033	439	.70±.023	442	.62±.026	554

* Final over-all key based on all samples.

statistical analysis indicates that, to a high degree, background items which predict social acceptance at one economic level will also predict at other economic levels.

To analyze further the characteristics of this common predictive background, the items were grouped by five experts in the field of adolescent behavior into areas descriptive of adolescent activities. Only items were included in this grouping which were significant ($p \leq .20$) in predicting the criterion in three or more of the six economic samples. The items included were further screened to make certain that they were significant in at least one sample of each economic level.

Descriptions from common predictive items follow by areas for the accepted and rejected boys.

Area I. Family and Home Experiences

Accepted boys: The parents, in the opinions of the accepted boys, possess the following characteristics: (*a*) They are pleased with the school marks their children obtain. (*b*) They do not show any favoritism in the family but treat all the children alike. (*c*)

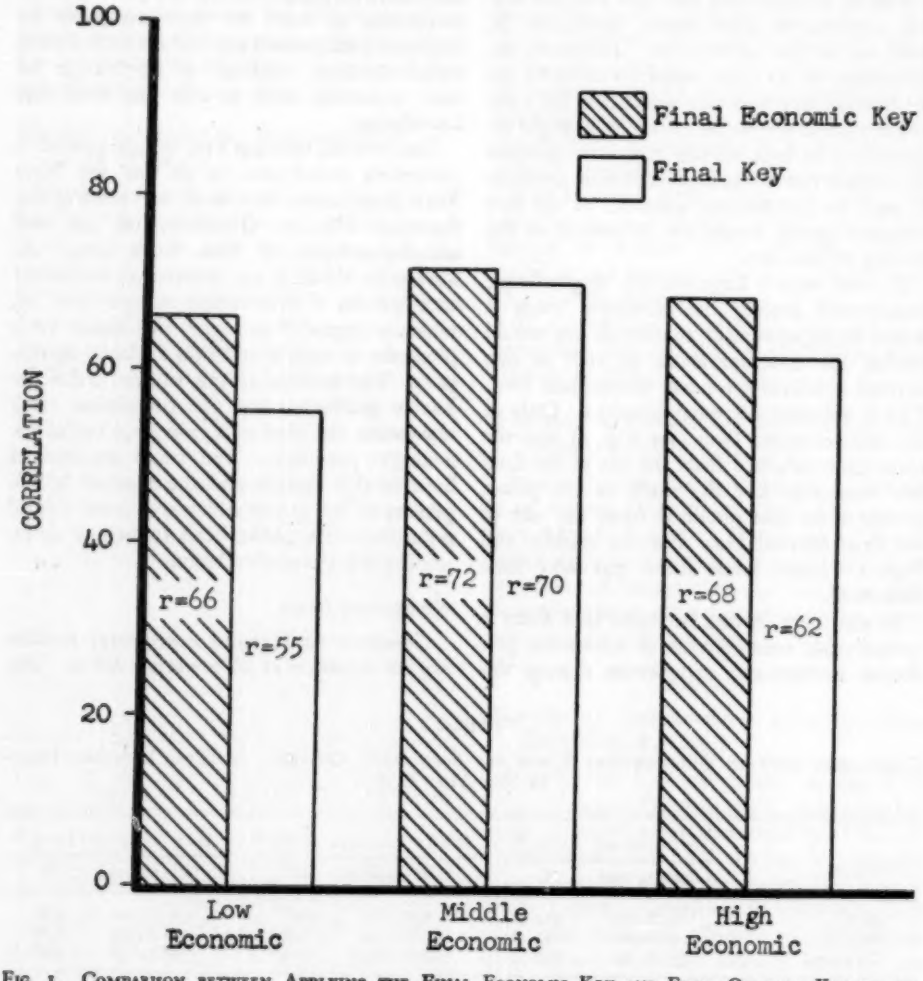

FIG. 1. COMPARISON BETWEEN APPLYING THE FINAL ECONOMIC KEY AND FINAL OVER-ALL KEY TO ALL SAMPLES

The fathers are interested in athletic activities such as golf, tennis, gymnasium.

Rejected boys: The parents, in the opinion of the rejected boys, are not satisfied with the grades of their children in school and are not interested in any social or athletic organizations.

Area II. Intellectual Activities and School Experiences

Accepted boys: The accepted adolescents at all economic levels can be described in the following manner: (*a*) Their marks were in the top 25 per cent of their class and they received prizes because of excellence in school grades. They felt that their academic achievements had been the result of hard work. (*b*) They received their best grades in arithmetic. (*c*) These boys have established good relationships with their teachers, most of whom they like and who they feel have had an influence on them, causing them, for example, to become very much interested in some particular subject.

Rejected boys: The rejected adolescents, in contrast, obtained grades in the third 25 per cent of their class and achieved their best marks in shop work. They disliked many of their teachers, who they felt had little influence over them.

Area III. Social Activities

Accepted boys: Sports and extracurricular activities took most of the accepted adolescents' time after school. These boys played regularly on football, baseball, and basketball teams, and they preferred the "prestige" positions, i.e., they wanted to be in the infield in baseball, and in football they desired to be the fullback, halfback, or end. They have also been captains of athletic teams, presidents of their class, and chairmen of committees.

Rejected boys: The rejected group, however, did not play regularly on teams either outside or within school. Their outside activities were usually hobbies, reading for pleasure, listening to the radio and television, or attending movies. These boys have never been elected to any school office or committee.

Area IV. Beginnings of Economic Independence

Rejected boys started working earlier (at the age of 11 or 12) than did their accepted peers, and they obtained most of their spending money from their own earnings. This is especially characteristic of rejected boys at the low and middle economic levels.

Area V: Adjustment to Concept of Self

Accepted boys: Accepted boys considered themselves most successful in getting along with their classmates, and they also felt that they make friends quickly and have many more close friends than most of the boys in their class. This is true at all economic levels, but particularly at the middle and high. In athletics, they felt that they can play most games better than the average person of their own age and sex.

Rejected boys: The rejected boys, on the other hand, had only one or two close friends and they considered themselves only average in ability to make new friends. They likewise considered their athletic ability "about average." The only two things they thought they could do better than the average person of their own age and sex were to "get along with their girl friends and to shoot a gun."

TABLE 3

CORRELATIONS BETWEEN BACKGROUND INFORMATION QUESTIONNAIRE WITH SOCIOMETRIC CRITERION SCORES IN AREAS OUTSIDE OF NEW YORK CITY

GEOGRAPHIC AREA	*N*	FINAL OVER-ALL QUESTIONNAIRE KEY BASED ON NEW YORK CITY POPULATIONS	
Westchester (New York State)	300	.48±	.045
Indiana	100	.25±	.096

Summary for All Activity Areas

Accepted boys generally. Accepted boys appeared to have made an adequate adjustment to their parents and teachers. Their parents are pleased with their grades and

they do not, in the opinion of the adolescents, have any favorites in the family group. The parents are interested in athletics, which is characteristic of the boys themselves who have played on many school teams. The boys' athletic and social pursuits did not interfere with their scholastic activities, for their grades were usually excellent. These boys believed they possess many more close friends than most of the boys in their class and they felt that they make friends more quickly. They appeared to have the ability to "relate" themselves to many individuals at the same time.

Rejected boys generally. In contrast, the rejected adolescents have more withdrawn personalities. Their parents are not interested in athletics or social pursuits or in their boys' athletic or social activities. The boys themselves do not play on teams in school and have only one or two close friends. They seem to prefer pursuits which can be conducted alone, such as hobbies, reading for pleasure, listening to the radio or attending movies. Their parents are dissatisfied with their school grades, which are usually in the lower 25 per cent of their class. The rejected boys fail to establish a good relationship with their teachers, most of whom they dislike. Further, these boys generally start to work at an early age, earning more money than the accepted boys.

Summary of Differences Among Economic Levels

Accepted boys. Close family relationships are predictive of accepted adolescents at the high and middle economic levels but not at the low.

At the middle economic level, with regard to school activities, these boys stated that their parents have had the greatest influence over them in the choice of school subjects. Also, they felt that English was the school subject that influenced them most. However, they spent the same amount of time studying each subject.

The accepted adolescents *at the high economic level* are granted more personal freedom than at the middle or low, i.e., they are allowed to spend more time away from home in socially approved activities, such as attending summer camps at least four times by the time they were twelve, and staying overnight at a friend's home. They also felt satisfied with their fathers' attitude toward them and they do not desire to do more for their children than their father did for them.

Rejected boys. Being an "only child" in a family is a predictor of rejection *at the high economic level* only. These rejected boys also felt that they would give their children more affection and consideration as individuals than their fathers gave to them.

The adolescents *at the middle economic level* whose mothers work in positions outside of the home are more likely to be rejected. They state that they have attended three different schools and have disliked six or more teachers since the first grade. They felt that shop work is the only course that really interests them. These adolescents claimed that they could play card games better than the average person of their own age.

At the low income level only, affiliation with the Boy Scouts is a predictor of rejection. These boys also preferred women teachers who have taken a great interest in their personal affairs.

Background experiences which indicate greater social and economic independence appear to characterize rejected boys at middle and low economic levels. These boys started to work at an early age and plan to attend a business school rather than a college. Further, they solve their personal problems without consulting their parents or teachers. At the middle level, the rejected boys have had their first evening date when under 14 years of age; and at the low economic level, they have spent a month away from home when under 11 years of age.

The reversals in the items, i.e., (alternatives predictive of accepted boys at one economic level and rejected at another) indicate that those children whose intellectual and social activities are patterned after a higher economic level than those of their classmates are likely to be rejected. For example, being interested in music (i.e., listening to concerts, participating in glee clubs), and reading five to nine magazines a week are characteristic of accepted adolescents at high or middle economic levels and rejected adolescents at the low economic level.

SUMMARY AND CONCLUSIONS

The problem in this study was to determine whether there are background experiences significantly related to social acceptance among adolescent boys and, if so, to ascertain whether these background experiences are characteristic generally of all adolescent boys or whether there are differences dependent upon their economic status.

Student choices of seat mates were used as the criterion of social acceptance, and estimated financial income of parents as the criterion of economic status. All students in the various samples completed a Personal History Questionnaire which was constructed to represent six areas of adolescent adjustment. Two samples were taken at each of three contrasting economic levels in New York City and two others were drawn at undifferentiated levels from other geographic areas for validation of results. The total population tested consisted of approximately 2,000 adolescent boys between the ages of 13 and 16 years. The chi-square statistic was utilized to ascertain the degree of significance of each item alternative in the questionnaire in predicting the criterion. It was then possible to determine what specific background experiences were predictive of social acceptance for a particular economic level and for adolescent boys generally. The predictive background items were combined into scoring keys, the Personal History Questionnaire scored, and correlations computed between these scores and the criterion scores. The correlation obtained with the economic scoring keys and the over-all scoring key were compared to indicate the degree of commonness among the economic levels in predictive background experience.

A majority of the items predicting social acceptance at one economic level will also predict social acceptance at other levels. Accepted boys generally appear to have made an adequate adjustment to their parents and teachers. Their parents are pleased with their school grades and they do not, in the opinion of the adolescents, have any favorites in the family group. They report that they are proficient in both athletic and scholastic activities and possess many more close friends than most of the boys in their class. Rejected boys have poor relations with teachers, are less proficient in athletic and scholastic activities and are more withdrawn, have fewer close friends.

There are some background items which predict in only one or two levels. For example, rejected adolescents at the high economic level, only, are more likely to be "only children," while background experiences which indicate greater economic independence (starting to work at an early age and working at more jobs) appear to characterize rejected boys at the middle and the low, but not at the high economic level.

Scores obtained by means of an over-all scoring key predicted the criterion at each economic level with but a slight loss, as compared with prediction from keys for the separate economic levels. When the final over-all key was applied to the three economic levels, the correlations were as follows: high, .62, middle, .70, low, .55. For samples from other regions of the country, correlations of .48 and .25 were obtained. The latter correlation was from a highly homogeneous group on which the criterion was inadequately established. These results are shown in Fig. 1 and Table 1.

1. Descriptive items of adolescent activities can be selected to predict social acceptance. It has been possible to score background experiences with reliability coefficients of better than .93 and validity coefficients of .62 or better.

2. In addition, the items predictive of social acceptance show a high degree of consistency between economic levels. It is true that there are differences between economic levels in background experiences; the influence of these differences upon the development of certain behavior patterns has been demonstrated by previous studies. However, as far as one can generalize from the present study, those background experiences predictive of social acceptance are, for the most part, common to all economic levels.

REFERENCES

1. BONNEY, M. E. The consistency of sociometric scores and their relationship to teachers' judgment of social success and to personality self-ratings. *Sociometry*, 1943, 6, 409–424.

2. BROWN, A. W., MORRISON, J., & COUCH, G. Influence of affectional family relationship on character development. *J. abnorm. soc. Psychol.*, 1947, 42, 422–428.

3. FRIEND, J. G., & HAGGARD, E. A. Work adjustment in relation to family background. *Appl. Psychol. Monogr.*, 1948, **16**, 135–150.
4. GUTHRIE, E. R. Personality in terms of associative learning. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. Vol. I. New York: Ronald, 1944. Pp. 49–68.
5. HAVIGHURST, R., & TABA, H. *Adolescent character and personality*. New York: Wiley, 1949.
6. HOLLINGSHEAD, A. B. *Elmstown's youth*. New York: Wiley, 1949.
7. JENNINGS, H. H. *Leadership and isolation*. New York: Longmans, Green, 1950.
8. JOHNSON, H. M. (In cooperation with BOOTS, M., & WHERRY, R.) *On the actual and potential value of biographical information as a means of predicting success in aeronautical training*. Washington: Civil Aeronautics Administration, 1942.
9. PARTRIDGE, E. DeA. *Social psychology of adolescence*. New York: Prentice-Hall, 1939.
10. REALS, W. H. Leadership in high school. *Sch. Rev.*, 1938, **46**, 523–531.
11. RIESMAN, D. *The lonely crowd*. New Haven: Yale Univer. Press, 1950.
12. RUNDQUIST, E. A. Personality tests and prediction. In D. H. Fryer and E. Henry (Eds.), *Handbook of applied psychology*. New York: Rinehart, 1950.
13. WATTENBERG, W. Delinquency and only children: study of a "category." *J. abnorm. soc. Psychol.*, 1949, **44**, 356–366.
14. *Construction and selection of items for the biographical information blank*. Washington: Adjutant General's Office, Personnel Research Section, War Department, Report No. 703, 1945.
15. *The history and development of the biographical inventory*. Washington: National Research Council Committee on Selection and Training of Aircraft Pilots, Division of Research, Report No. 70, 1946.
16. *Validation of a program for the selection of officers for retention in the peacetime Army*. Washington: Adjutant General's Office, Personnel Research Section, War Department, Report No. 704, 1946.

Received April 26, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

IS THERE A MECHANISM OF PERCEPTUAL DEFENSE?

LEO POSTMAN, WANDA C. BRONSON, AND GEORGE L. GROPPER
University of California

SEVERAL investigators have reported higher thresholds of perceptual recognition for negatively valued or taboo materials than for neutral or positively valued materials (2, 3, 6, 8, 11, 14, 17). *Perceptual defense,* an unconscious mechanism of resistance to recognition of threatening stimuli, has frequently been invoked to explain the differences in speed of recognition. Is such a concept necessary?

Whatever its face validity, it is uneconomical to consider perceptual defense as a special mechanism of *perception* as long as the experimental findings can be subsumed under more general principles. Such reanalysis would deny neither the importance of motivational factors in perception nor their relevance to personality theory. The aim is to formulate the determinants of perceptual behavior at the highest level of generality possible. Many determinants are equally applicable to meaningless and meaningful, need-relevant, and neutral materials. The contributions of specific motivational variables must be evaluated against a base line specified in terms of such general determinants.

A principle of perceptual defense was first considered by Bruner and Postman (2, 11) in explanation of high thresholds for negatively valued and emotionally charged words. The individual tends to ward off the recognition of threatening and inimical stimuli as long as possible within the constraints of the stimulus situation. Although this statement is descriptively true for some situations, it is now clear that the original formulation of the principle of perceptual defense was inadequate because (*a*) it did not specify the mechanisms producing the raised thresholds and (*b*) it failed to consider explanation of the observed facts in terms of more general determinants of perceptual recognition. In a later theoretical statement (9) one of us suggested, for example, that apparently defensive thresholds may be due to the competition of alternative recognition responses (perceptual hypotheses). Interference by competing responses is, of course, a general principle not restricted to emotionally charged materials.

Perceptual defense, as a *special perceptual mechanism,* was affirmed and elaborated in a study by McGinnies (6). In his experiment, McGinnies compared recognition thresholds for emotionally toned taboo words and neutral words. Without exception, thresholds were higher for emotional words than for neutral words. In addition, the *Ss*' GSR's were measured during the period preceding recognition in an attempt to obtain evidence for systematic reaction to the stimulus words prior to correct verbal report. GSR's prior to recognition of the critical words were found to be significantly greater than those preceding recognition of the neutral words. Finally, *Ss*' incorrect guesses suggested a tendency actively to avoid recognition of the emotionally toned words. McGinnies considered these findings as strong evidence for the operation of perceptual defense and went on to speculate concerning the neurological mechanisms mediating the unconscious defensive activity.

McGinnies' study precipitated severe criticisms of perceptual defense as an unwarranted *ad hoc* assumption. These criticisms, advanced by Howes and Solomon (4), directly lead to new and critical questions which must be answered experimentally in order to clarify the status of perceptual defense. There are, in the opinion of Howes and Solomon, two more plausible, and simpler, reasons for McGinnies' results. First and foremost, the taboo words used by McGinnies are, according to the Thorndike-Lorge word count (16), much less familiar words than the neutral words with which they were compared. On the basis of the known general relationship between word familiarity and recognition speed (5, 13), lower thresholds for the neutral words are to be expected regardless of differences in emotionality. Second, Howes and Solomon consider it very probable that McGinnies' *Ss* tended to withhold overt re-

ports of the socially unacceptable words until they were completely sure of having recognized them correctly. McGinnies, however, has defended his original interpretation (7, 8). He is not convinced that the lower thresholds for neutral words can be ascribed to their greater familiarity. The correlation between familiarity and speed of recognition "does not constitute other than circumstantial evidence for the assumption that the higher thresholds of the taboo words were determined by their alleged infrequency of occurrence. If . . . the elevated duration thresholds for these words were a function of their affective connotation, the net regression effect would be the same" (7, p. 237). We can decide between the alternative explanations only by comparing thresholds for neutral and taboo words which have been equated for familiarity.

There is no substantial evidence, according to McGinnies, that *S*s do, in fact, withhold their verbal reports of taboo stimulus words. As a positive indication against the hypothesis of selective verbal report, he presents the finding that recognition thresholds for *neutral* words are raised when these words are exposed immediately following the perception of taboo words (8). These results, he argues, constitute indirect evidence for perceptual defense since the avoidance reaction appears to generalize to temporally adjacent neutral words which the *S* ordinarily has no reluctance to report. We do not believe that these findings represent cogent evidence for perceptual defense or dispose of the problem of selective report. Perceptual defense has been alleged to be an avoidance reaction specific to threatening stimuli and must be demonstrated *as such* before it is appropriate to assert its generalization to neutral stimuli. By the same token, one would not speak of the generalization of a conditioned response before the presence of such a response was established. It is clear that we must attack the problem of selective verbal report directly in order to gauge its contribution to relative thresholds for neutral and taboo words.

Finally, McGinnies feels that the systematic differences between prerecognition responses to neutral and emotional words constitute further evidence for perceptual defense (cf. also 11). He classified the prerecognition responses into four categories: (*a*) *structurally similar*, i.e., words resembling the stimulus words in letter structure; (*b*) *structurally unlike;* (*c*) *nonsense*, i.e., words without dictionary meaning; and (*d*) *part*, i.e., disconnected groups of letters. Proportionately more similar and part responses were given to the neutral words, and proportionately more unlike and nonsense responses to the critical words. McGinnies interprets the greater frequency of unlike and nonsense hypotheses as "tactics apparently designed to delay accurate recognition of the stimulus word" (6, p. 250). Such responses, he argues, represent distortions by which *S* can avoid seeing the dangerous stimulus. It is difficult to insist on this interpretation because exactly the opposite findings could be equally well reconciled with the hypothesis of perceptual defense. If there were more structurally similar and part responses to the charged words, one could argue that perceptual defense delays final correct organization of the word even though the conditions are adequate for recognition of component parts of the stimulus. Structurally similar and part hypotheses could with great facility be labeled compromises between the impact of the stimulus and *S*'s "resistance to its recognition!" The process of perceptual defense has not been sufficiently specified to make the reported differences in prerecognition responses relevant to its demonstration. Such differences can equally well be ascribed to the uncontrolled variations in the familiarity of the words.

Other recent experiments which have been offered in support of perceptual defense are beset by similar difficulties. Rosenstock (14) compared recognition thresholds for neutral and affectively charged sentences and found the latter subject to perceptual "repression." The only control of familiarity consisted of matching the second clauses of critical and neutral sentences in terms of the "number of meaning units." Clearly *number* of words, syllables, etc. does not provide an estimate of the familiarity of a sentence or part of a sentence. Whether and to what extent the thresholds for the two types of sentences were a function of differential emotional content must remain indeterminate. Eriksen's findings (3) concerning recognition thresholds for neutral and need-relevant pic-

tures are equally difficult to interpret. When need-relevant pictures are recognized more slowly than neutral pictures we cannot speak of defense (or of relative amounts of defense in the case of different pictures) unless and until we have shown that both sets of stimuli have been equated in terms of structural characteristics and familiarity.[1] It should be noted that both Rosenstock and Eriksen instructed their *Ss* to expect unusual stimuli in order to decrease the likelihood of selective verbal report. The effects of such instructions can be fully evaluated, however, only by comparison with the effects of other instructions.

The problems raised by these experiments have important general implications. They bring us squarely up against the question whether what appears to be perceptual defense can be reduced to the operation of determinants which are not specifically emotional. The controversy can be resolved only by direct experimental confrontation of the alternative explanations. If the difference between thresholds for emotional and neutral words can be varied at will by manipulating factors other than the emotional significance of the stimulus words, the status of perceptual defense will be rendered highly doubtful.

METHOD

Major variables. This experiment represents an extension and modification of McGinnies' study on perceptual defense. In the light of the criticisms advanced against McGinnies' findings and interpretations, the effects of two variables on recognition thresholds for neutral and taboo words were investigated: (*a*) the relative familiarity of the two classes of stimuli and (*b*) conditions affecting *Ss*' set and readiness to give verbal reports of taboo words. The hypothesis was tested that it was the failure to take these variables into account which was largely responsible for the threshold differences observed by McGinnies. It follows from this hypothesis that the difference between thresholds can be increased or decreased in predictable ways by proper manipulation of the variables of familiarity and set.

The factor of familiarity. Control of familiarity was achieved by the following procedure. Like McGinnies, we used seven taboo words and eleven neutral words. The taboo words were the same as those used by McGinnies. Each of the taboo words was matched in familiarity with a neutral control word by means of the Thorndike-Lorge semantic count (16). The remaining four neutral words had frequencies of usage corresponding to the average frequency of the taboo words. Two of the taboo words, which do not appear in the Thorndike-Lorge semantic count, were matched with control words having a frequency of one (i.e., one in 4½ million words) according to the semantic count.[2] The list of stimulus words and their frequency values appear in Table 1.

The question may be raised whether equating of familiarity in terms of the Thorndike-Lorge word count is appropriate. This count is based on the frequency with which words appear in a broad sample of literary sources. Taboo words probably appear less frequently in written English than they do in conversation, and hence the word count may systematically underestimate their frequency of usage. It should be remembered, however, that the tachistoscopic presentation of words (used by both McGinnies and ourselves) creates essentially a *reading* situation for the *Ss*. An estimate of frequency based on usage in reading material is, therefore, reasonable. On the other hand, it is not known to what extent visual duration thresholds are influenced by the frequency of spoken as well as written usage. The possibility must be acknowledged, therefore, that we have un-

TABLE 1

LIST OF STIMULUS WORDS

Words are listed in the order of presentation. Taboo words are in italics.

WORD	FREQUENCY VALUE *
lotus	10
capon	10
raped	13
lathe	11
belly	47
noted	45
cleat	1
whore	4
tiara	4
kotex	1**
mixer	1
terse	4
penis	1**
hilly	10
knack	13
filth	4
clove	4
bitch	4
Av. taboo	10.6
Av. neutral	10.3

* Frequency of occurrence in 4½ million words according to the Thorndike-Lorge semantic count.

** Estimated frequency.

[1] These criticisms do not affect the validity of differences among groups of *Ss* exposed to the same sets of materials. The interpretation of these differences as reflecting various degrees of *defense* is, however, questioned.

[2] Such an estimate is appropriate in view of the probability (discussed below) that the frequency of taboo words is systematically underestimated by the Thorndike-Lorge count.

derestimated the familiarity of the taboo words. Such an error would not, however, vitiate the test of the experimental hypothesis. According to McGinnies, the difference between thresholds for neutral and taboo words is due to a defensive reaction and is not primarily determined by their relative familiarity. Thresholds for taboo words should be *higher* even if the control of familiarity is short of perfect. Application of the familiarity hypothesis, on the other hand, would predict *lower* thresholds for taboo words if their frequency had been systematically underestimated, or *equal* thresholds if the estimate were correct.

The factor of selective verbal report. If *Ss* deliberately withhold their reports of taboo words, the thresholds for such words would appear to be unduly high. The possibility that *Ss*' reports are deliberately suppressed or falsified exists in every perceptual experiment and can never be completely discarded. In the case of recognition of taboo words, this problem may be critical since *Ss* have to give verbal responses which they have presumably been taught to avoid. It is futile to argue after the fact whether or not there was such deliberate suppression or falsification. We can, however, attempt to manipulate explicitly the conditions which we believe contribute to *Ss*' readiness to report faithfully what they perceive. The resulting changes in perceptual response can then provide us with an estimate of the amount of variance attributable to *Ss*' readiness to report.

In this experiment, manipulation of *Ss*' readiness to report taboo words was attempted in two ways: (*a*) by preliminary instructions and (*b*) by varying the *S-E* relationship.

Four different sets of instructions were used. The *Uninformed Group* was given the usual instructions for tachistoscopic recognition, without any indication that taboo words would be included among the stimuli:

> This is an experiment on visual perception. We want to find out how well you can recognize words that are presented for short periods of time. When I say "Ready," fixate your eyes on this screen. About one second later a word will be flashed for a brief period of time and will appear on the screen. As soon as the word has been flashed on, write down everything you saw or thought you saw, including single letters, on the record sheet.

The *Informed Group* was given the same general instructions, with the following additional statement informing them of the presence of taboo words among the stimuli:

> Some of the words which will be shown are of the kind not commonly used in polite society, especially in the presence of members of the opposite sex. Remember, some, not all, of the words that we will show are of this type.

The instructions given to the *Facilitation Group* were designed to discourage *Ss* from withholding reports of taboo words. These *Ss* were given the same instructions as the Informed Group with the following additional statement:

> It has been found that some people find it difficult to recognize such words when they are presented for brief periods of time and are anxious about reporting them when they have seen them. In general, difficulties in seeing and reporting such words are a sign of emotional trouble. Some types of patients show this difficulty to a severe degree.

Finally, the *Inhibition Group* was given instructions designed to maximize the tendency to withhold reports of taboo words. These *Ss* were given the same instructions as the Informed Group but with this additional statement:

> It has been found that some people find it difficult to recognize such words when they are presented for brief periods of time and are anxious about reporting them when they have seen them. In general, difficulties in seeing and reporting such words are a sign of good mental adjustment. Most normal people show this kind of reaction to some extent, particularly people who are socially successful.

To the extent that these different instructions were effective in manipulating *Ss*' set and readiness to report taboo words, we would expect the measured thresholds for taboo words to be (*a*) lower for the Informed Group than the Uninformed Group, and (*b*) lower for the Facilitation Group than the Informed Group, since the former is both informed about the presence of taboo words *and* encouraged to report them. It would be difficult to predict exactly where the Inhibition Group should fall since it has the advantage of expecting the appearance of taboo words but is discouraged from reporting them. In any event, the thresholds of the Inhibition Group for taboo words should be higher than those of the Facilitation Group.

An additional experimental treatment was introduced in a further attempt to manipulate *Ss*' readiness to report. Most of the taboo words are of a kind that one would be particularly reluctant to use in the presence of the opposite sex. Deliberate withholding of taboo words might be expected to occur more often in the presence of an *E* of the opposite sex than in the presence of an *E* of the same sex. Two *Es*—one male and one female—were, therefore, used in the experiment. Under each of the four major experimental conditions, each *E* tested two groups of *Ss*, one group of males and one group of females. There were, therefore, 16 different experimental subgroups. Each of the subgroups consisted of 6 *Ss*, giving a total of 96 *Ss*.

Subjects. The *Ss* were undergraduate students at the University of California who had volunteered for the experiment. They were assigned to the various experimental conditions at random. The *Ss* did not know the purpose of the experiment.[8]

Apparatus and procedure. The stimulus words, typed in capital letters, were shown on 2 × 2 lantern slides. The slides were exposed on a ground-

[8] Members of the Facilitation Group and Inhibition Group were, of course, disabused concerning the instructions at the end of the experimental session but admonished not to discuss the experiment with others.

glass screen from a projector equipped with photographic shutter. The speed of exposure was held constant at .01 sec. The threshold was approached by increasing the brightness of the exposure flash in even steps. Each successive exposure represented an increase of 1.3 volts. Each slide was exposed 18 times, which was sufficient in virtually all cases for correct recognition of the word. The *Ss* were tested in groups of three. They wrote their responses on prearranged record sheets, using a separate sheet for each word.

RESULTS

Analysis of recognition thresholds. The number of exposures required for correct recognition of a word was used as a measure of the threshold. As usual, *Ss*' tachistoscopic acuity varied widely. In order to make the relative thresholds for neutral and emotional words independent of *S*'s acuity level, each *S*'s threshold measures were converted into standard scores. All tabulations and comparisons will be presented in terms of standard scores. Table 2 presents the summary of an analysis of variance performed on the transformed threshold data. We shall refer to the results of this analysis throughout our discussion of the experimental results.

Average thresholds for neutral and emotional words. Table 3 shows the average recognition thresholds for neutral and emotional words for the four major experimental groups and for all groups combined. Under all conditions, the thresholds for taboo words are *lower* than the thresholds for neutral control words. This difference is highly significant. Clearly, there is no evidence here for perceptual defense. Taboo words are recognized sooner than neutral words, probably because of a systematic underestimation of the familiarity of the taboo words. When neutral words no longer have the advantage of greater familiarity, the raised thresholds for taboo words disappear. In fact, the *direction* of the difference in recognition thresholds has been changed by manipulation of the familiarity variable.[4]

TABLE 2

SUMMARY OF ANALYSIS OF VARIANCE OF RECOGNITION THRESHOLDS

SOURCE	*df*	SUM OF SQUARES	MEAN SUM OF SQUARES	*F*	*p*
Stimulus Words	1	598.55	598.55	98.12	<.01
Sex of *Ss*	1	1.17	1.17	—	—
Sex of *E*	1	0.42	0.42	—	—
Conditions	3	4.31	1.44	—	—
Stimulus Words × Conditions	3	48.67	16.22	2.66	.05
Stimulus Words × Sex of *Ss*	1	30.88	30.88	5.06	.01–.05
Stimulus Words × Sex of *E*	1	2.76	2.76	—	—
Conditions × Sex of *Ss*	3	6.22	2.07	—	—
Conditions × Sex of *E*	3	4.06	1.35	—	—
Sex of *E* × Sex of *Ss*	1	2.30	2.30	—	—
Stimulus Words × Conditions × Sex of *Ss*	3	20.03	6.67	1.09	—
Stimulus Words × Conditions × Sex of *E*	3	62.23	20.74	3.40	.01–.05
Stimulus Words × Sex of *E* × Sex of *Ss*	1	17.58	17.58	2.88	—
Conditions × Sex of *E* × Sex of *Ss*	3	14.28	4.76	—	—
Pooled Error*	163	994.62	6.10		
Total	191	1808.08			

* Individual differences among *Ss* treated alike and the third-order interaction, which were not significant, were included in the error term.

The effects of preparatory instructions. As Table 3 and Fig. 1 show, the observed relative thresholds for neutral and emotional words vary with the instructions given at

[4] Our attention has just been called to a recent communication by McGinnies and Adornetto in the newsletter *Cognitive Processes* (compiled by the Department of Psychology, University of Alabama). They report an experiment in which a group of normals and a group of schizophrenics were tested with taboo words (the same as used previously) and neutral words equated for familiarity. They still find higher thresholds for taboo words although the difference is considerably smaller than in McGinnies' original study. For the normal *Ss* the difference now is only 8.2 milliseconds as compared with the previously reported difference of 45 milliseconds. The remaining discrepancy between this result and ours may be due, of course, to the different neutral words used. At best, then, the raised thresholds are a "now-you-see-it-now-you-don't" phenomenon once the familiarity factor has been equated.

TABLE 3

AVERAGE RECOGNITION THRESHOLDS (STANDARD SCORES) FOR TABOO AND NEUTRAL WORDS UNDER THE DIFFERENT EXPERIMENTAL CONDITIONS

GROUP	TABOO WORDS	NEUTRAL WORDS
Uninformed	—.11	.10
Informed	—.22	.14
Facilitation	—.29	.20
Inhibition	—.20	.14
All Groups	—.20	.14

words once they had been recognized. The Facilitation Group, which had been explicitly discouraged from withholding reports of taboo words, has the lowest relative thresholds for such words. This group probably had lower criteria of certainty for reporting emotional words than any other group. The Inhibition Group, which was encouraged to hesitate about reporting taboo words, has higher relative thresholds for emotional words than the Facilitation Group. The results for the Inhibition Group are virtually identical

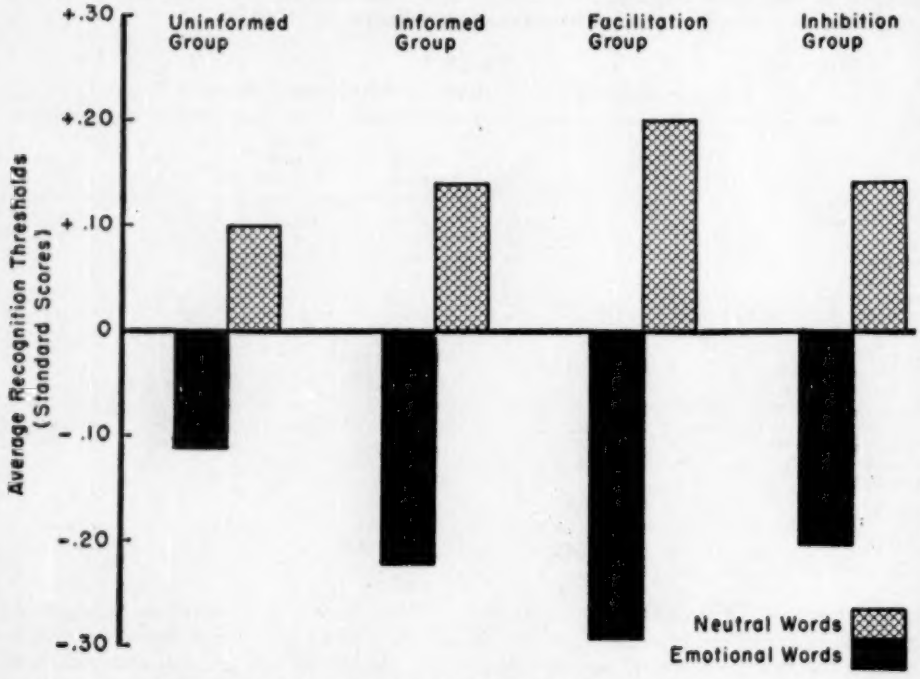

FIG. 1. AVERAGED RECOGNITION THRESHOLDS FOR TABOO AND NEUTRAL WORDS UNDER THE DIFFERENT EXPERIMENTAL CONDITIONS

the beginning of the experiment. The interaction, Conditions × Stimulus Words, is statistically significant. The Uninformed Group has higher relative thresholds for taboo words than any of the groups that had been forewarned about the nature of the stimuli. We may assume that the warning to expect taboo words produced a preparatory set which facilitated recognition of such words. In addition, however, the variations among the forewarned groups suggest that the instructions affected *Ss*' readiness to report taboo with those in the Informed Group. In the absence of any explicit statement concerning speed of reporting, the *Ss* in the Informed Group may have adopted the same cautious criteria for verbal report as did the *Ss* in the Inhibition Group. It appears, then, that preparatory set and selective verbal report contribute to the observed threshold differences.

Sex differences. Male *Ss* recognize taboo words more readily than female *Ss* (Table 4). This difference is significant. The thresh-

olds of both men and women *Ss* are, however, equally affected by the nature of the preliminary instructions: the interaction, Conditions × Stimulus Words × Sex of *Ss*, is not significant.

Experimenter-subject relationship. Table 4 fails to reveal any tendency for slower recognition of emotional words in the presence of an *E* of the opposite sex than in the presence of an *E* of the same sex. If anything, the opposite seems to be the case. The interaction, Stimulus Words × Sex of *E* × Sex of *Ss*, is, however, not significant.

TABLE 4

AVERAGE RECOGNITION THRESHOLDS (STANDARD SCORES) FOR TABOO AND NEUTRAL WORDS, GROUPED BY SEX OF *E* AND SEX OF *Ss*

SEX OF *E*	MALE *Ss*		FEMALE *Ss*	
	TABOO WORDS	NEUTRAL WORDS	TABOO WORDS	NEUTRAL WORDS
Male	—.22	.16	—.22	.15
Female	—.28	.20	—.10	.08
Av.	—.25	.18	—.16	.12

Other interactions. As Table 2 shows, there is only one other interaction which approaches significance: Stimulus Words × Conditions × Sex of *E*. Inspection of the data indicates that the female *E* obtained larger systematic threshold differences from condition to condition than did the male *E*. All *Ss*, male and female, tended to conform more to the instructions in the presence of the female *E*. We do not know, of course, whether this result is a function of the particular *Es* in this experiment or whether it has wider generality. In any event, this finding confirms the view that the reporting situation may have significant effects on the observed recognition thresholds.

Prerecognition responses. As we pointed out above, we do not consider the classification of prerecognition responses used by McGinnies crucial to the demonstration of perceptual defense. It is of interest to point out, however, that we were unable to duplicate his findings concerning the differential distributions of prerecognition responses for emotional and neutral words. Table 5 shows the percentages of prerecognition responses for the two types of words falling into the various content categories used by McGinnies. In order to make our findings as comparable as possible with his, Table 5 presents only the findings for our Uninformed Group, which was given the same type of instructions as were his *Ss*. (The distributions of prerecognition responses are, however, substantially the same for all our groups.) Clearly, there is no systematic difference between neutral and emotional words with respect to the distributions of prerecognition responses. Our neutral words were, of course, different from those used by McGinnies. With familiarity held constant, the differences in prerecognition responses, like the raised thresholds, are absent.

TABLE 5

PERCENTAGES OF PRERECOGNITION RESPONSES FALLING INTO VARIOUS CONTENT CATEGORIES

TYPE OF RESPONSE	TABOO WORDS	NEUTRAL WORDS
Similar	18.30	19.96
Part	67.12	64.73
Unlike	3.05	2.32
Nonsense	11.52	12.98

DISCUSSION

Familiarity, motivation, and recognition thresholds. In McGinnies' study, the observed differences in recognition thresholds for neutral and emotional words could be ascribed either to differences in familiarity or to the operation of perceptual defense. In our experiment, we have attempted to "untie" these two factors. With familiarity controlled, there is no indication of perceptual defense. In fact, the emotional words have lower thresholds, probably because of an underestimation of their familiarity.

One may object at this point that the effects of familiarity and perceptual defense are not mutually exclusive. Is it not possible that the relative thresholds for the taboo words would have been even lower had it not been for the operation of perceptual defense? If the familiarity of the emotional words has, in fact, been underestimated, this error may obscure the effects of perceptual defense. In advancing such an argument, one would deprive the concept of perceptual defense of most of its substance. One would concede its effectiveness to be so small as to be counteracted by an error in the estimate of familiarity.

In any event, the invocation of an hypothetical effect which is obscured by the operation of a known variable would carry little conviction.

We are emphatically not arguing here that recognition thresholds are completely determined by the sheer frequency with which the subject has been exposed to, and/or has used, the stimulus words. Frequency is a major parameter of whatever processes produce variations in perceptual readiness, whether they be S-R connections, "hypotheses," or what have you. There is some evidence that motivational factors interact with frequency of occurrence to determine readiness for perceptual recognition of words (10, 12, 13). It appears, however, that the taboo characteristics of the words used in this experiment do not produce effects which interact significantly with familiarity. In any event, whatever evidence *is* found for the effectiveness of motivational determinants does not call for a proliferation of special perceptual mechanisms, but is most economically analyzed in the framework of a general theory of perception which provides for the conceptualization of motivational variables (1, 9).[5]

Selective verbal report and recognition thresholds. The significant variations in thresholds produced by differences in instructions can be interpreted in two alternative ways. We may assume that preliminary information concerning the presence of taboo words (*a*) aroused perceptual expectancies which facilitated recognition of such words (cf. 12, 15, 18), and (*b*) influenced *Ss*' readiness to report such words. Such an interpretation appears to be in agreement with the general trend of the experimental results. An alternative formulation might assert that the nature of the instructions influenced the degree of perceptual defense. The variations in thresholds produced by instructions may be regarded as evidence of different strengths of "resistance" to the recognition of taboo words. We reject this alternative explanation for the following reasons. First, since perceptual defense has not been demonstrated in terms of the over-all trends in thresholds, it would be inappropriate to speak of variations in degree of defensive response induced by instructions. Furthermore, and this is the theoretically more important consideration, the processes described by such terms as "defense" or "repression" surely are not to be conceived as subject to significant manipulation by means of preparatory instructions! If they were so manipulable, such defenses or repressions would, indeed, be very different from those described by personality theorists. We prefer, therefore, to think of the instructions as producing variations in *Ss*' set and selective verbal reports.

Attempts to control *Ss*' set may, of course, fail. Our manipulation of the *E-S* relationship did not produce consistent results. (It is possible that this procedure would have been more effective if *Ss* had been required to report their responses orally rather than to write them down.) Whatever trends did appear, e.g., the tendency toward greater conformity with instructions in the presence of the female *E*, emphasize the need to consider the influence of *situational* variables on perceptual thresholds.

Sex differences. The fact that women had significantly higher relative thresholds for emotional words than men is again open to alternative interpretations. "Sex difference" is not a unitary psychological variable. We can only guess as to the characteristics of male and female *Ss* which result in the threshold differences. Consistently with our general argument, we may assume that female *Ss* are (*a*) less familiar than the male *Ss* with the taboo words used in this experiment, and (*b*) are less ready to report these words, i.e., may require a somewhat higher level of certainty for reporting.

One may, on the other hand, assert that female *Ss* are more prone to perceptual defense than are male *Ss*. Again, we would object to such an explanation because it is inconsistent with the major experimental results. We see, moreover, no cogent theoretical reasons for postulating the type of personality differences between male and female *Ss* which such an interpretation would require.

The status of perceptual defense. We conclude that the experimental findings to date have failed to lend support to the concept of perceptual defense. The results have been

[5] McGinnies' recent attempt (8) to translate defense into S-R reinforcement terms represents a move in this direction.

either indeterminate or can be explained in terms of more general principles. It may, of course, be argued that stimuli such as those used by McGinnies and ourselves are not appropriate for testing the hypothesis of perceptual defense. Such may or may not be the case; the possibility of future positive evidence cannot, of course, be excluded. At the present, however, perceptual defense has, at best, the status of an unconfirmed hypothesis.

SUMMARY

The concept of *perceptual defense* has been used to explain raised recognition thresholds for emotional stimuli as compared with neutral stimuli. The invocation of such a principle would, however, appear unnecessary if the facts can be subsumed under more general principles of perceptual behavior. This experiment was designed to answer the question whether there is any evidence for perceptual defense in word recognition when the factors of familiarity, set, and selective verbal report are taken into account.

A series composed of taboo words and neutral words was presented tachistoscopically for recognition. The familiarity of the two types of words was equated as closely as possible by means of the Thorndike-Lorge word count. Four different sets of instructions were used. The Uninformed Group was not given any indication that taboo words would be included among the stimuli. The Informed Group was explicitly warned to expect taboo words. The Facilitation Group was given instructions designed to discourage *Ss* from withholding reports of emotionally charged words. Finally, the instructions given to the Inhibition Group were intended to maximize *Ss*' hesitation about the reporting of taboo words. Two different *Es*, one male and one female, conducted the experiment. Under each of the four major experimental conditions, each *E* tested both a group of male *Ss* and a group of female *Ss*.

Analysis of the recognition thresholds yielded the following results:

1. Under all conditions of the experiment, the thresholds for taboo words were somewhat *lower* than the thresholds for neutral words. The difference in thresholds is probably due to a systematic underestimation of the familiarity of the taboo words.

2. The relative thresholds for neutral and emotional words varied significantly with the nature of the preliminary instructions. The Uninformed Group had higher relative thresholds for taboo words than any of the groups forewarned to expect taboo words. The Facilitation Group had the lowest relative thresholds for such words. The results for the Inhibition Group are almost identical with those for the Informed Group. These findings indicate that set and selective verbal report contribute to the observed threshold differences.

3. Women have higher relative thresholds for taboo words than do men. It is possible that women are less familiar with these words than are men and/or that women are more prone to selective verbal report.

4. Relative thresholds for taboo words are not significantly different in the presence of an *E* of the opposite sex than they are in the presence of an *E* of the same sex. There is, however, a significant tendency for all *Ss* to conform to instructions more fully in the presence of the female *E*. The *E-S* relationship has demonstrable effects on the observed threshold differences.

The experimental results thus fail to provide any support for a mechanism of perceptual defense. The discussion stresses the inadequacy of *ad hoc* mechanisms for the analysis of motivational factors in perception. The effects of such factors are most economically treated in the framework of a general theory of perception.

REFERENCES

1. BRUNER, J. S. Personality dynamics and the process of perceiving. In R. R. Blake and G. V. Ramsey (Eds.), *Perception: an approach to personality*. New York: Ronald, 1951.
2. BRUNER, J. S., & POSTMAN, L. Emotional selectivity in perception and reaction. *J. Pers.*, 1947, **16**, 69–77.
3. ERIKSEN, C. W. Perceptual defense as a function of unacceptable needs. *J. abnorm. soc. Psychol.*, 1951, **46**, 557–564.
4. HOWES, D. H., & SOLOMON, R. L. A note on McGinnies' "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, **57**, 229–234.
5. HOWES, D. H., & SOLOMON, R. L. Visual duration thresholds as a function of word probability. *J. exp. Psychol.*, 1951, **41**, 401–410.
6. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.

7. MCGINNIES, E. Discussion of Howes' and Solomon's note on "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, 57, 235–240.
8. MCGINNIES, E., & SHERMAN, H. Generalization of perceptual defense. *J. abnorm. soc. Psychol.*, 1952, 47, 81–85.
9. POSTMAN, L. Toward a general theory of cognition. In J. H. Rohrer and M. Sherif (Eds.), *Social psychology at the crossroads.* New York: Harper, 1951.
10. POSTMAN, L., & BROWN, D. R. The perceptual consequences of success and failure. *J. abnorm. soc. Psychol.*, 1952, 47, 213–222.
11. POSTMAN, L., BRUNER, J. S., & MCGINNIES, E. M. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1946, 43, 142–154.
12. POSTMAN, L., & LEYTHAM, G. Perceptual selectivity and ambivalence of stimuli. *J. Pers.*, 1951, 19, 390–405.
13. POSTMAN, L., & SCHNEIDER, B. H. Personal values, visual recognition and recall. *Psychol. Rev.*, 1951, 58, 271–284.
14. ROSENSTOCK, I. M. Perceptual aspects of repression. *J. abnorm. soc. Psychol.*, 1951, 46, 304–315.
15. SIIPOLA, E. M. A group study of some effects of preparatory set. In J. J. Gibson (Ed.), Studies in psychology from Smith College. *Psychol. Monogr.*, 1935, 46, No. 6 (Whole No. 210), 27–38.
16. THORNDIKE, E. L., & LORGE, I. *The teacher's word book of 30,000 words.* New York: Teachers College, Columbia Univer., 1944.
17. VANDERPLAS, J. M., & BLAKE, R. R. Selective sensitization in auditory perception. *J. Pers.*, 1949, 18, 252–266.
18. ZANGWILL, O. L. A study of the significance of attitude in recognition. *Brit. J. Psychol.*, 1937, 28, 12–17.

Received May 1, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

A CROSS-CULTURAL INVESTIGATION OF CLOSURE[1]

DONALD N. MICHAEL
Boston University

ACCORDING to Gestalt theory, to quote Köhler, "sensory organization . . . appears as a primary fact which arises from the elementary dynamics of the nervous system" (4, p. 199). To be sure, the Gestaltists have indicated that perceptual organization is the product of both internal and external organizing forces: those "within the process in distribution itself and which will tend to impress on this distribution the simplest possible shape" and those "between this distribution and the stimulus pattern which constrain the stress toward simplification" (5, p. 138). But while the Gestaltists concede the effects of external stimulus constraints on the organization of perceptions, it seems safe to say that they have not considered in detail the effects of previous experience and learned responses on the so-called internal organizing forces. That is, given the same external stimuli, will *S*s with different previous experience with those stimuli organize their perceptions similarly? The problem is not whether perceptual organization is or is not completely independent of neural processes, since obviously it must be dependent on them at some level. Rather, the problem is how far along a perceptual organization continuum, so to speak, do such processes determine perception independently of the particular experience context of the individual.[2] It may be that some of the so-called "sensory organization . . . which arises from the elementary dynamics of the nervous system" are instead universally *learned* types of sensory organization within a given culture. If this is so, then it is possible that such learned responses are masked by their cultural universality and thereby are taken for innate. Therefore, it was the purpose of this study to determine whether the process of closure is such a learned response or whether the Gestalt contention that closure is a general law of innate perceptual organization tends to be substantiated.

In order to determine whether or not the tendency to closure is dependent on culturally determined perceptual organization, it was necessary to compare response patterns of two cultures, one of which stressed closure and one of which stressed nonclosure, both in terms of the actual stimulus patterns (e.g., ornamental designs) prevalent in the culture and in terms of the values (i.e., directives for behavior) stressed in the cultures. The two cultures used in this experiment were Anglo-Saxon American and Navaho Indian.

It seems safe to propose that Western European culture in general and American culture in particular stress the concept of closure. In the area of asthetics and design, unbroken continuity (e.g., streamlining) and symmetry are emphasized. And among the ideal behavior patterns favored are those stressing task completion, both materially and temporarily. (Consider, for example, the Zygarnic effect.)

On the other hand, students of the Navaho have long remarked on the Navaho "fear of completing anything" (7, p. 226). "As a spirit outlet the basketmaker leaves an opening in the design; the weaver leaves a small slit between the threads . . ." (7, p. 226) and so on. In Navaho values there is a "need to have no completely enclosing frame around any of the works so that the evil inside can have an opening through which to leave" (8, p. 12). In discussing the dimension of generality in value systems, Kluckhohn says: "a negative value in Navaho culture is fear of closure" (6, p. 413). Not only is emphasis on nonclosure primary in the Navaho value system, but, as indicated above, it pervades

[1] This research was facilitated by the Laboratory of Social Relations, Harvard University, in conjunction with the research project on the Comparative Study of Values. The research was undertaken with funds provided by the Social Science Division of the Rockefeller Foundation. The writer wishes to thank Professor Jerome Bruner for suggestions on the design of this experiment and Professors Clyde Kluckhohn and Evon Vogt for advice and assistance in the field. He also wishes to thank Dr. Eleanor Maccoby and Professor Nathan Maccoby for suggestions on the analysis and interpretation of these data.

[2] See, e.g., Bruner, J. S. Personality dynamics and the process of perceiving. In G. Ramsey and R. Blake (Eds.), *Perception: an approach to personality*. New York: Ronald Press, 1951. Ch. V.

their designs. For example, on pottery an ornamental band around the pot is left incomplete. There is an opening on the order of 1 to 5 degrees in the band. Thus, while the Navaho exposure to circles is by no means limited to those with openings, it is probable that they are systematically exposed to more experiences stressing nonclosure than would be ascribed to white Americans.

It was presumed that if Navaho experiences of the sort described above contribute to a differential capacity, compared to whites, to respond to nonclosure there should be unacculturated Navaho Indians who reflect this differential capacity.

The stimuli selected for testing the *Ss*' responses were a series of circles and double concentric circles with and without breaks in their periphery. Circles were selected as the stimuli because the Gestalt experiments on closure had been performed using circles and because the circle is a familiar visual pattern in the Navaho culture. Furthermore, in many of its occurrences in Navaho culture, the circle is deliberately not completed, i.e., a small opening is left in the periphery. The double concentric circle was also utilized because it is similar to both a common pattern used in Navaho sandpaintings and to many design patterns in white American culture.

Therefore, the null hypothesis to be tested was: there is no significant difference in the response behavior (i.e., tendency to closure) to a set of open and closed circles between samples of white Americans and Navaho Indians.

METHOD

Subjects

The sample of 20 white Americans consisted of male and female members from the very small farming community of Homestead, New Mexico, located in the same region as that from which the Navaho sample was obtained.[3] These *Ss* were exposed to the stimuli in a darkened warehouse.

The Navaho sample of 20 *Ss* consisted of male and female members of the surrounding Navaho reservation.[4] They were exposed to their stimuli in a darkened building attached to a local trading store.

Apparatus

The experimental situation consisted of exposing the *Ss* to a series of circles, with varying degrees of incompleteness, for .1 second. The stimuli were projected on 2×2″ slides through a translucent plastic screen (the screen being between the projector and the subjects) and appeared on the screen as 12 inches in diameter.[5] The *Ss* sat four to six feet from the screen. Thus the circle subtended a visual angle of 7.1 to 5.6 degrees.

The size of incompletion of the stimulus circle ranged from zero to eight degrees in one-degree steps. A double concentric circle with a break in the outer circle of 0, 1, or 7 degrees was also used. Further, it was possible to control the light intensity of the projected image (the Variac was patched into the light circuit) and the Ilex No. 3—Acme shutter screwed to the front of the projection lens allowed variations in time of exposure. The use made of these extra facilities and additional stimuli will be discussed later.

Procedure

Each *S* had a pencil, pad of paper, and writing board on which to record what each thought he or she saw.[6] There was sufficient light projected through the screen between stimuli projections to make this procedure entirely feasible. The *Ss* were treated in groups of from 3 to 6 persons per group.

Directions were given orally by the experimenter to the white group and orally through a Navaho interpreter to the Navaho group.[7] The instructions were as follows:

> I want to know what you see on my screen when I show a picture to you for a very little time. The pictures are simple designs. Each time I show a picture you draw what you see on the screen on a page of the pad of paper you have. Draw one picture on each page, using one page after the other. If you don't see anything on the screen, leave that page empty and turn over to the next page. Because I want to know what you yourself first see, please don't look at what the other people are drawing and don't talk about it. I will tell you when I am going to show each picture. Then I will show the picture. Then you draw what you saw. Then I will tell you to turn the page. Then I will

[3] Various research projects involving this community are still under way. It is therefore necessary to preserve the anonymity of the community by giving it a fictitious name in order that there be no possibility that information of the results of research on the community can affect its attitude toward the present researchers, thereby distorting future research findings.

[4] It is the judgment of Professors Evon Z. Vogt and Clyde Kluckhohn that the Navahos used in this experiment can be considered as unacculturated representatives.

[5] The projector was a 35-mm commercial model manufactured by the Three Dimensional Corporation of Chicago, Illinois under the designation, "TDC Vivid." The 300-watt projection lamp was driven by 110 volts at 60 cycles and blower cooled. Light intensity was adjusted through a Variac (General Radio Co. type V-5MT) set at a scale reading of 40.

[6] It should be noted that all Navahos are able to use a pencil and paper as means for representing what they see.

[7] Professor Clyde Kluckhohn checked the Navaho interpretation of the directions and was of the opinion that it adequately expressed the purposes and directives of the English version.

tell you that I am going to show another picture, and so on. Are there any questions?. . . Now we will start. Here is a picture.

After the recording was completed each pad of pictures was checked while *S* was still present so that any drawings which might be misinterpreted could be discussed with him. Finally the *Ss* were requested not to discuss their experience with others because it would spoil them for the experiment.[8]

RESULTS

The four measures used for determining differential capacity to perceive the openings in the circles were: (*a*) the total number of openings seen; (*b*) the number of persons seeing the openings regardless of size; (*c*) the mean number of openings seen per person; and (*d*) the median size of openings seen.

TABLE 1

RESPONSES TO CIRCLE STIMULI

SUBJECTS	NUMBER OF CIRCLES SEEN			NUMBER OF PERSONS SEEING ANY CIRCLES		
	OPEN	CLOSED	TOTALS	OPEN	CLOSED	TOTALS
Navaho	43	137	180	16	4	20
White	55	125	180	18	2	20
Totals	98	262	360	34	6	40

The distribution of the number of circles seen as open and as closed by the whites and the Navahos appears in Table 1. Note that since, conceivably, the truly closed circle could have been seen as open, responses to it are included in this distribution. A $p=.20$, obtained by Fisher's exact method for chi square (3, pp. 96–97), indicates that there is no significant difference between the total number of openings seen by the whites and Navahos.

Reference to Table 2 gives the number of persons seeing *an* opening. Here $p=.66$ as obtained by Fisher's exact method for chi square, indicating no significant difference between the total number of whites and Navahos seeing any opening at all.

The mean number of openings seen per person for the Navahos was 2.69; for the whites it was 3.06. The standard t test gives a $t=.555$ which is a p of approximately .5. This too is clearly not a significant difference. Further, since the data are considerably skewed, and since one would expect such a distribution of the mean number of openings to be skewed toward the lower end, it was felt advisable to use Festinger's F test (2). Here an F of 1.22 was obtained which is less than the F of 1.83 required for significance at the .05 level.

The final test for significance was on the median size openings seen by whites and Navahos. These medians may be regarded as a measure of that size opening for which the probabilities of a "closed" response and an "open" response are equal. The median size opening for Navahos was 6 degrees and for the whites 5.4 degrees. This gives a $t=.718$ or a p of approximately .45. As such, there is no significant difference between the median size opening seen by whites and Navahos.

Thus the null hypothesis to the effect that there are no significant differences in the tendencies of whites and Navahos to perceive open circles as closed is not disproved.

DISCUSSION

The fact that this is a cross-cultural study requires that several other conditions must be considered which might conceivably lead to the same results as those indicated above, but for reasons having little to do with the presence or absence of innate neurological processes such as closure. Some of these factors which possibly are culturally controlled follow.

[8] Because of the nature of the relationship between the researchers and the natives, both white and Indian, which has been developed over many years, it is most likely that this request was carefully honored. At any rate, all *Ss* were asked preliminary to the experiment if they knew what it was about, and none indicated that they did.

Visual Acuity

If the Navaho's ability to make fine visual discriminations is less than that of the white sample, then we would hardly expect them to see openings when they are physically incapable of doing so under the best of seeing conditions. The possibility that Navahos simply cannot see as well as whites was controlled as follows. In all cases where *S* recorded no openings during the exposure, a circle with a 1° opening was projected for an unlimited time and *S* was then asked again to report on what he saw. In the very few cases where *S* was unable to report the opening the data were discarded. (This procedure was also used for the white sample.)

TABLE 2

NUMBER OF PEOPLE SEEING 7° OPENING IN CIRCLE AT VARIOUS LIGHT INTENSITIES

LIGHT INTENSITY*	NUMBER OF PEOPLE SEEING OPENING		CUMULATIVE NUMBER OF PEOPLE SEEING OPENING	
	WHITE	NAVAHO	WHITE	NAVAHO
15	1	1	1	1
20	5	1	5	2
25	6	2	7	4
30	10	2	11	5
35	14	8	15	9
40	12	8	15	11
45	11	9	16	11
50	12	7	16	12

* The numbers indicating the light intensity settings are the same as those on the indicator dial on the Variac. The larger the number, the brighter the image.

Light Intensity

Do Navahos require more light to discriminate as well as whites? Since *E* had control over light intensity, some pretesting was done in order to find an intensity level "acceptable" to both whites and Navahos. Setting the Variac at a scale value of 40 seemed to be a good working value. However, each *S* was also exposed to a 7° opening at a variety of light intensities so that more precise comparisons could be made later. The data are given in Table 2.

While there is a difference in the number of *S*s seeing an opening of 7° at a light setting of 40, the difference is not significant, either in terms of the actual percentage of the population seeing the opening or the cumulative percentage of the population seeing the opening. The *p*'s are .34 and .36, respectively. They are based on the difference between proportions, corrected for continuity, between uncorrelated samples (1, pp. 76–83). It is true that for values rather lower than 40, the whites consistently see better than the Navahos (except at the lowest intensity of 15), but it is clear that at the light intensity used in obtaining the results reported above, no significant difference existed in the tendency of either group to see the opening. Further, essentially the same results obtain when the double concentric circle with the 7° opening in the outer one is presented under varying light intensities (see Table 3). Thus, at the light intensity used, there is no reason to believe that the Navahos cannot discriminate as well as do the whites.

TABLE 3

NUMBER OF PEOPLE SEEING 7° OPENING IN DOUBLE CONCENTRIC CIRCLE COMPARED TO LIGHT INTENSITY

LIGHT INTENSITY*	NUMBER OF PEOPLE SEEING OPENING		CUMULATIVE NUMBER OF PEOPLE SEEING OPENING	
	WHITE	NAVAHO	WHITE	NAVAHO
15	0	2	0	2
20	3	2	3	3
25	10	3	10	5
30	10	2	10	5
40	11	8	13	9
50	9	10	13	11

* See footnote under Table 2.

"Set" Tendencies

The Navaho tendency to "sit tight" has often been commented on by anthropologists. "When in a new and dangerous situation do nothing. If a threat is not to be dealt with by ritual canons, it is safest to remain inactive. If a Navaho finds himself in a secular situation where custom does not tell him how to behave, he is usually ill at ease and worried. . . . The American tradition says: 'When danger threatens, do something.' The Navaho tradition says: 'Sit tight and perhaps in that way you may escape evil' " (7, p. 226). Is this tendency operating here to a greater extent than it is on the white sample? If it

is assumed, for example, that the initial exposure is seen as a closed circle, will there be a greater tendency on the part of the Navaho to be set to continue to see them as closed? In an effort to test for "set," *E* did the following experiment. One set of *Ss* was exposed to circles with openings of 8°, 5°, and 1°, exposed in sequence from the largest to the smallest, and each of the circles was exposed for 5, 1, 1/2, 1/5, 1/10, and 1/25 of a second in that order. Each opening was presented through all time exposures before the next size opening was presented. Another set of *Ss* was exposed to precisely the same variables, only in their case the order was ascending: i.e., the smallest opening was exposed first and for the shortest time interval. Unfortunately, the only available white and Navaho samples were too small to be statistically stable. However, analysis of the data would seem to indicate that, trendwise at least, perceptual set tendencies for this sort of stimulus are not importantly different between the two groups.

Unfamiliarity with the Experimental Situation

Certainly one feels that the total experimental situation is likely to be less familiar to the Navaho than to the whites even if the whites do live in a small farming community. As such, one might expect this to have some effects on Navaho behavior. However, it is difficult to see how unfamiliarity can account for tendencies to see figures as closed or as open unless one hypothesizes which type of response is more "natural" in an unfamiliar situation. But it is the problem of determining, so to speak, which type of response is the more natural that is the subject of this research. Thus, while unfamiliarity may affect the perception of open or closed circles, we cannot say which way the effect would work without begging the question.

Effect of Context

It is possible that, because the experimental context is probably relatively meaningless for the Navaho, openings were not perceived whereas *in* context they would have been perceived. In other words, fewer breaks are seen precisely because, unless the context is important, no careful visual discrimination of the breaks occurs: i.e., everything that is not explicitly (contextually) broken is *ipso facto* closed.[9] This would be in keeping with Kluckhohn's description of the Navaho characteristic of describing experience in terms of events rather than by attributes, especially when it is remembered that nonclosure in Navaho culture usually occurs in specific contexts (e.g., sand paintings, pottery, and basket ornamentation).

It is difficult to verify this point experimentally. For one thing, there was no way of determining whether a nonclosure response, in a context, was by association with the rest of the context or independent of the context. Or to state the point in more Gestaltist terms, it would be difficult in complex stimulus situations to predict the relative degree of dominance and interaction of various tendencies to pragnanz. Moreover, there do not seem to be any contextually equivalent white and Navaho stimulus situations.[10]

Thus within the range of considerations discussed above, the results of this experiment do not disprove the null hypothesis which was that there is no significant difference in the response behavior (i.e., tendency to closure) to a set of open and closed circles between samples of white Americans and Navaho Indians. This finding does not refute the Gestalt contention that closure is an innate and not a learned process. It does not prove, however, that closure is not a learned process since it is still possible that closure is a consequence of culture-bound

[9] Note, however, that if one accepts this argument one also accepts the implication that the capacity to discriminate breaks as such is learned and that nonperception of breaks is not due to autochthonous processes which, as an argument in this experimental situation, seems to beg the question.

[10] In an effort to expose the Navaho to a stimulus of at least potential contextual significance, the 20 white and 20 Navaho subjects used in the main experiment described above were also exposed to a double concentric circle with an opening of 0, 1, or 7 degrees in the outer one. This design abstractly corresponds to the basic design of many sand paintings. Using the same analysis procedures described above, the results were: (*a*) for the total number of openings seen, $p=.52$; (*b*) for the total number of persons seeing the openings, $p=.74$; (*c*) for the mean number of openings seen per person, $t=.82$ or $p=.4$. Because only three different sized openings were used, no useful estimate of the median size opening is available.

However, the three measures reported indicate clearly no significant difference in behavior of whites and Navahos to this stimulus.

learning which is not detectable by the methods used herein. The results of this research would be critical only if there were differences in the data. Then it could be stated that the ability to perceive closure is learned.

It may thus be argued that these results are, possibly, consequences of unspecified or indeterminate cultural characteristics which may override any "real" differences in perceptual response, learned or innate. That such may indeed be the case argues well for more cross-cultural research. Until more such research is undertaken, there will be considerable indeterminacy about the contributions to behavior, in the laboratory situation, of innate or pan-cultural processes, compared to the contributions of culture-bound learning, which may underlie the whole repertory of behavior that the *S* (and *E*) brings to the laboratory situation.

SUMMARY

The purpose of this experiment was to determine the effect of differences in cultural conditioning on the perception of closure. One sample of *Ss* was Navaho Indians living in a culture stressing nonclosure, both in values and in ornamental design. The other sample of *Ss* was white Americans who stressed closure both in values and design. Closure was investigated by determining the differential capacity of the two samples of *Ss* to detect small openings in tachistoscopically presented circular stimuli. If the perception of closure is innate, as proposed by the Gestaltists, then culturally determined learning should have no effect on the preception of closure. If closure is primarily a learned perception, the differences in cultural experiences of the two samples might result in differences in ability to detect closure. The *Ss* were presented with circles with different size openings for various lengths of time. The measures used for determining differential capacity to perceive the openings in the circles were (*a*) the total number of openings seen; (*b*) the number of persons seeing the openings regardless of size; (*c*) the mean number of openings seen per person; and (*d*) the median size of openings seen. The results, for all conditions, gave *p*'s which indicated no significant differences in the perception of closure between the two cultures. A number of factors involving culture differences which may have contributed to the results are discussed. The point is emphasized that considerable cross-cultural research will be necessary in order to eliminate the indeterminateness of these factors.

REFERENCES

1. EDWARDS, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
2. FESTINGER, L. Statistical test for means of samples from skew populations. *Psychometrika,* 1942, 8, 205–210.
3. FISHER, R. A. *Statistical methods for research workers.* New York: Hafner, 1948.
4. KÖHLER, W. *Gestalt psychology.* New York: Liveright, 1947.
5. KOFFKA, K. *Principles of Gestalt psychology.* New York: Harcourt, Brace, 1935.
6. KLUCKHOHN, C. Values and value orientation in the theory of action. In T. Parsons & E. Shils (Eds.), *Toward a general theory of action.* Cambridge: Harvard Univer. Press, 1951. Pp. 388–433.
7. KLUCKHOHN, C., & LEIGHTON, DORTHEA. *The Navaho.* Cambridge: Harvard Univer. Press, 1946.
8. LEE, DORTHEA D. *Religious perspectives of college teaching in anthropology.* New Haven: The Edward W. Hazen Foundation, 1951.

Received May 1, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

SHARING LEADERSHIP IN SMALL, DECISION-MAKING GROUPS [1]

LEONARD BERKOWITZ

Combat Crew Training Research Laboratory
Human Resources Research Center

SOCIAL scientists engaged in research in the area of leadership have approached the exceedingly complex problems in this area from many different directions. The personalities of the leaders (2) and of the followers (12) have been investigated, as have the behavioral characteristics of elected or sociometrically chosen leaders (8) and of institutionalized or designated leaders (13). But despite the approach used, as Katz indicates (9), the present trend is to emphasize the relational aspects between the leader and his group. The characteristics of the leader, whether personality-wise or behavioral, become significant only in terms of the leader's group. A leader's behavior, thus, may or may not satisfy the needs of the group, and a group member's behavior may or may not be in conformity with the traditions of his group.

In this sense, training in "democratic" leadership may not benefit the leaders of certain groups, particularly if those groups have learned to expect and adjust to, or even require, more "authoritarian" leadership practices. "Democratic" leadership may affect the group adversely (for example, in terms of criteria such as satisfaction with the group meeting) because it is not in keeping with the group's needs and expectations. The present study does not test the efficacy of "democratic" groups as contrasted with "authoritarian" groups. However, one aspect of "democracy" in groups may be described in terms of leadership sharing. We may define a "democratic" group as one in which many members may influence the group in the course of its goal setting and goal achievement. French has pointed out (5) that many conditions may intervene to affect the relationship between group "democracy" and criteria of group effectiveness. One such condition may have to do with the expectations of the groups.

Very broadly, then, this is the problem with which we are presently concerned. What happens when the functions of the institutionalized or designated leader are shared by others in the group?

METHOD

The present study is based on the methodology and many of the hypotheses developed by the Conference Research project. The most extensive study undertaken by the project was a field observation of 72 conferences in government and industry. The present report is concerned with a small part of the data collected in the course of this study.

The Data

Of the 72 conference groups in the sample, 42 were industrial in nature, i.e., they were conference groups in organizations devoted to manufacturing. Fourteen were business groups from organizations engaged in retailing, banking, newspaper publishing, and hotel enterprises. The remaining 16 came from federal and local governmental organizations. The size of the groups ranged from 5 to 17 members with a mean of 9 members, and the duration of the conferences was from 16 to 191 minutes. Many of these groups, furthermore, were composed of members having high status in their organizations. Of particular relevance to the present study is the chairman's status relative to the mean status of the group as measured by distance on the organizational chart from the head of the organization. Only 9 of the chairmen are at or below the mean level of the group; the great majority of the 72 chairmen are thus above the mean of the group in organizational status. Lastly, the sample of conferences was restricted to those primarily decision-making in nature.

Methods of Data Collection

Three observers, each with a specific area of interest, were present at all the meetings. The independent measures used in the present study were largely obtained from the records of one of these observers. The task of this individual, the "prob-

[1] This study is part of a larger series of investigations undertaken by the Conference Research project at the University of Michigan. Thanks are due to the project staff, and to Drs. R. W. Heyns and H. Guetzkow in particular, without whom this study would not have been possible, and to Dr. Daniel Katz and the members of my doctoral committee for their helpful suggestions. An overview of some of the general findings of the project is presented elsewhere (11). The full data upon which this report is based are to be found in the author's dissertation, Some Effects of Leadership Sharing in Small, Decision-Making Conference Groups, 1951, University of Michigan Library, and in the reports of the Conference Research project (1).

lem-solving coder," was to code the remarks of the participants according to the problem-solving category system developed by Heyns (7). In essence, the observer noted every participation—the total speech of an individual from the time he started to speak until the next person spoke—breaking each down into the categories it included.

Upon the completion of the meeting, each conference member was given a five-item questionnaire attempting to assess his satisfaction with various aspects of the meeting. Some of these ratings, all highly intercorrelated, were later averaged to form a measure of the group's satisfaction with the conference. After these ratings were collected—but before they were examined—the observers left the meeting room and immediately made a series of 54 ratings describing the interactions among the conference participants. The final source of data came from interviews with the participants within 48 hours after the meeting. Among other things, the data gathered at this time had to do with the members' perceptions concerning the interactions within the group.

The Measures

In general, the functions or behaviors going to make up the role of the conference chairmen are behaviors either directly influencing the group to follow him or behaviors facilitating this influence. Previous exploratory work carried out by the project led to the development of hypotheses as to what some of these behaviors might be, and measures of these were used as the independent variables in the present study. The meaning of these variables is twofold. In a sense they are two sides of the same coin. The variables have to do with the extent to which conference leadership is shared by members other than the designated leader, the chairman. When the chairman is said to be firmly in control over the group's procedure, it is also implied that there are few other group members attempting to influence the course of the meeting. Again, the greater the proportion of solution proposing contributed by the other conference members the smaller the proportion of all the solution proposing accounted for by the chairman. The independent variables are as follows:

1. *Leader Control of Process*—postmeeting observer rating. This item refers to the leader's control over the *manner* in which the group went about handling its problems, and not the content of the problems.

2. *Leader Permissiveness in Content*—postmeeting observer rating. This measure is a rating of the designated leader's restrictiveness in the area of the *content* of the group's problems.

3. *Functional Differentiation of the Leader*—postmeeting observer rating. This has to do with whether or not the conference chairman behaved differently from the other group members by performing somewhat unique functions. A separate analysis has revealed that there were generally two or more behavioral leaders among the group members when the chairman was rated low on this item. Thus, the more the designated leader was differentiated from the others, the more the leadership functions were centered in him alone.

4. *Percentage of Member Participation*—observer tally made during the meeting. This is the percentage of all the participation in the group that was accounted for by the members other than the chairman.

5. *Percentage of Functional Units to the Leader*—observer tally made during the meeting. Within each participation a group member may perform several different functions: he may propose a solution to the group's problems (solution proposing), develop and enlarge upon the solution (development giving), and seek further comments from the others (development seeking). Furthermore, each of these functional units may be directed to a different member. This, then, is the proportion of the total number of functional units contributed by the group members which were directed to the chairman.

The following 3 variables have to do with the functional units or categories within the participations. In computing the following scores we used the proportion of all the units within any one category that were accounted for by the members other than the chairman.

6. *Percentage of Goal Setting by Members.*
7. *Percentage of Solution Proposing by Members.*
8. *Percentage of Summarizing by Members.*

Three of the dependent measures used in the Conference Research field study are employed here. They are designed to assess some of the more important outcomes of conferences: the cohesiveness of the group, the members' satisfaction with their meeting, and the productivity of the meeting.

1. *Cohesiveness.* The series of experiments carried out by the Research Center for Group Dynamics (3, 4) have demonstrated the theoretical importance of this construct. The present definition is the sum of five highly interrelated observer ratings of the attractiveness of the group situation, e.g., the pleasantness of the group atmosphere, the degree of personal liking, etc.

2. *Satisfaction with the Conference.* This is a mean of five highly interrelated ratings made by the conference participants at the conclusion of their meeting or shortly thereafter. These include: satisfaction with the decisions arrived at, satisfaction with the leader, satisfaction with the group's process, etc.

3. *Productivity.* The operational definition of productivity is the proportion of agenda items completed of those brought up for consideration. This information was obtained from the tallies made by one of the observers in the course of the meeting.

RESULTS

Relationships Over All Groups in the Sample

Table 1 reveals the correlations between the measures of leadership sharing and the three criteria. The results with the group's satisfaction with its meeting are fairly clean

cut. The more the chairman is the sole major behavioral leader the more satisfied the group with its conference. This can be seen in the following correlations. Satisfaction increases: the more the chairman controls the group's process, the more functionally differentiated he is, and the greater the proportion of functional units addressed to him. Satisfaction also increased the less the members participate relative to all the participating done in the group, and the less they do of all the solution proposing done in the group.

But although the conference groups in the present sample tend to be less satisfied when leadership is shared—when others beeffects of situational conditions upon the above relationships. For example, they suggested that there are "negative reactions"—lowered cohesiveness and satisfaction—to leadership sharing in both more and less permissively led groups.

Although there is more sharing of the leadership functions in the more permissive situations, the chairman's control of process, the extent to which he is functionally differentiated from the others in the group and the percentage of units addressed to him tend to be positively correlated (statistically significant or almost so) with cohesiveness and member satisfaction in both the more

TABLE 1

CORRELATIONS BETWEEN CRITERIA AND MEASURES OF LEADERSHIP SHARING FOR ALL 72 GROUPS

MEASURE	COHESIVENESS	SATISFACTION WITH CONFERENCE	PRODUCTIVITY
Leader control of procedure	.04	.29*	—.09
Leader permissiveness in content	.57**	—.04	.05
Functional differentiation of leader	.15	.33**	—.04
% Participation by members	—.17	—.38**	—.20
% Functional units to leader	.10	.31**	.13
% Goal setting by members	—.02	—.23	.06
% Solution proposing by members	.01	—.30**	—.15
% Summarizing by members	—.24	—.18	—.03

* Significant at the 5% level of confidence.
** Significant at the 1% level of confidence.

sides the chairman perform leadership functions—it appears that productivity does not suffer. The indices of leadership sharing bear no relation to the proportion of agenda items completed of those considered.

The only leadership measure significantly related to the ratings of cohesiveness was the rating of the leader's permissiveness in the area of the content of the group's problems. The more permissive the leader the higher the rating of group cohesiveness. Just as good a case may be made for saying that permissiveness produces cohesiveness as for saying that cohesiveness gives rise to permissiveness. It is likely that both are correct, that allowing the group freedom to express opinions can both result from and enhance the attractiveness of the group situation.

The Effects of Situational Factors

It is not possible to go into all of the additional "breakdowns" carried out to test the permissively led and the less permissively led groups. Similarly, the proportion of all the participations in the group accounted for by members is related to decreased cohesiveness and satisfaction under both types of conditions. This is shown in Table 2.

It should be pointed out that though the results indicate that the present groups want behavioral control by the socially recognized conference leader, they do not indicate that they want substantive control over what should or should not be said.

An attempt was made to determine whether the negative effects of sharing could be attributed to those groups having an opposing relationship with the chairman. That is, is the lessened satisfaction due to the fact that the sharing may be done in opposition to the chairman?

The 72 groups in the study were subdivided into three categories: (*a*) Those groups having no consistent leadership sharers, as indi-

TABLE 2

CORRELATIONS BETWEEN CRITERIA AND MEASURES OF LEADERSHIP SHARING FOR GROUPS CONTRASTED ON THE CHAIRMAN'S PERMISSIVENESS IN THE CONTENT OF THE GROUP'S PROBLEMS

MEASURE	N		COHESIVENESS		SATIS. WITH CONF.		PRODUCTIVITY	
	HI PERM.	LO PERM.	HI PERM.	LO PERM.	HI PERM.	LO PERM.	HI PERM.	LO PERM.
Leader control of process	39	33	.38*	.21	.44**	.18	.08	—.22
Functional differentiation of leader	39	33	.37*	.24	.49**	.18	.14	—.18
% Participation by members	39	33	—.43**	—.29	—.35*	—.46**	—.33	—.25
% Functional units to leader	39	33	.27	.22	.27	.37*	.15	.17
% Goal setting by members	37	32	—.14	—.16	—.14	—.38*	.12	—.05
% Solution proposing by members	39	33	—.39*	.02	—.18	—.47**	—.17	—.23
% Summarizing by members	30	23	—.06	—.30	.03	—.38	.11	—.14

* Significant at the 5% level of confidence. ** Significant at the 1% level of confidence.

cated by a high rating on the continuum of functional differentiation of the chairman. There is evidence indicating that there were more emergent leaders the lower the rating of designated leader differentiation. (*b*) Those groups having a somewhat lower rating on the differentiation of the chairman scale, but who appeared to be in a positive, supporting relationship with the chairman. These are termed the *positive-sharing* groups. (*c*) Those groups having consistent leadership sharers, but who were in less of a supporting relationship with the chairman. These groups are termed the *negative-sharing* groups.

The positive-sharing groups were differentiated from the negative-sharing groups on the basis of the ratio of supporting to supporting-plus-opposing interchanges with the chairman. It was reasoned that this proportion of supporting remarks to and from the chairman could serve as an indication of the extent of the group's supporting relationship with the chairman. Those groups having a higher proportion of supporting interchanges are more likely to be in a positive, supporting relationship with the chairman.

Several conclusions may be drawn from the correlations shown in Table 3. We can organize them in terms of the relationships with the criteria. But what seems to be generally indicated by the results is that there is little evidence that the type of relationship with the chairman is of primary significance in determining the correlation between the measures of sharing and the criteria.

1. Cohesiveness.

a. The leader's permissiveness in the content of the group's problems is positively related to the rating of group cohesion in the less supporting, negative-sharing group as well as in the more supporting, no-sharing and positive-sharing groups.

b. Despite the general supporting relationship with the chairman in the positive-sharing groups, the group attractiveness tends to decrease the more the members participate relative to the chairman and the less they address their remarks to the chairman.

2. Satisfaction with the conference.

a. The greater the proportion of the total number of participations accounted for by the members other than the chairman, and the less clearly the chairman is the major be-

havioral leader by being functionally differentiated from the rest of the group the less the group's satisfaction in the more supporting positive-sharing groups.

In general, then, the indications are that cohesiveness and member satisfaction may be lessened by leadership sharing even though the group is in a supporting relationship with its designated leader, the chairman.

groups with urgent problems as being more motivated, and with this motivation focused on the reaching of definite problem decisions. There is also a tendency for these groups to have greater interdependence among the members.

The correlations shown in Table 4 fairly clearly indicate differences in relationships with the sharing measures in the more and

TABLE 3

CORRELATIONS BETWEEN THE CRITERIA AND MEASURES OF LEADERSHIP SHARING FOR THE THREE KINDS OF SHARING GROUPS

MEASURE	KIND OF GROUP	COHESIVENESS	SATISFACTION WITH THE CONFERENCE	PRODUCTIVITY
Leader control of process	o sharing	.02	.12	−.13
	+ sharing	−.10	.32	.17
	− sharing	.31	.22	−.10
Leader permissiveness in content	o sharing	.64**	.03	.11
	+ sharing	.43	−.12	.02
	− sharing	.59**	.01	−.06
Functional differentiation of leader	o sharing	−.36	.13	−.08
	+ sharing	.29	.54*	.41
	− sharing	.52*	.20	−.06
% Participation by members	o sharing	−.14	−.15	−.15
	+ sharing	−.54**	−.47*	−.42
	− sharing	−.24	−.59**	−.33
% Functional Units to leader	o sharing	−.13	.12	.01
	+ sharing	.50*	.31	.30
	− sharing	.22	.58**	.52*
% Goal setting by members	o sharing	.14	−.05	−.32
	+ sharing	−.25	−.15	.24
	− sharing	−.13	−.31	−.18
% Solution proposing by members	o sharing	.03	−.45*	−.35
	+ sharing	−.32	−.04	−.09
	− sharing	.12	−.25	−.31
% Summarizing by members	o sharing	−.22	−.36	−.09
	+ sharing	.02	−.16	−.15
	− sharing	−.48*	−.02	−.23

* Significant at the 5% level of confidence.
** Significant at the 1% level of confidence.

Correlations with the Sharing Measures in Groups Contrasted on the Urgency of Their Problems.

Unlike the correlations in the groups contrasted on the permissiveness of the leader, and the relationship between the group and its leader, there is evidence that the urgency of the group's problems may affect the reactions to the sharing of leadership. The groups in the present sample were "broken down" into groups with more and less urgent problems, depending upon the mean of the members' ratings on this scale. Judging from the correlations (not given here) with the urgency scale, we might characterize the less urgent situations. For one thing, the results suggest that the designated leader's control over the group and his being functionally differentiated from the rest of the group are related to group cohesiveness and satisfaction only when the problems confronting the group are not pressing ones. These correlations are significantly different from the *r*'s in the more urgent situations. The percentage of all the participations accounted for by the members is negatively related to cohesiveness and satisfaction in only the less urgent situations, but the differences in correlation are not significant.

It is also interesting to note that the rela-

TABLE 4

CORRELATIONS BETWEEN THE CRITERIA AND MEASURES OF LEADERSHIP SHARING IN GROUPS CONTRASTED ON THE URGENCY OF THEIR PROBLEMS

MEASURE	N		COHESIVENESS		SATIS. WITH CONF.		PRODUCTIVITY	
	HI URG.	LO URG.	HI URG.	LO URG.	HI URG.	LO URG.	HI URG.	LO URG.
Leader control of process	35	37	—.27	.34	—.04	.59**	—.17	—.23
Leader permissiveness in content	35	37	.79**†	.26	.12	—.19	.17	.05
Functional differentiation of leader	35	37	.02	.32	.06	.68**†	—.16	.09
% Participation by members	35	37	—.01	—.34*	—.26	—.51**	—.05	—.30
% Functional units to leader	35	37	.06	.15	.31	.32	.09	.17
% Goal setting by members	33	36	—.00	—.02	—.02	—.22	.02	.11
% Solution proposing by members	35	37	.30	—.37*†	—.17	—.39*	—.00	—.18
% Summarizing by members	27	27	—.26	—.23	—.18	—.36	—.20	.11

* Significant at the 5% level of confidence. ** Significant at the 1% level of confidence.
† Correlation significantly different from other correlation at the 1% level of confidence.

tionships with the percentage of all the solution proposing accounted for by the members vary depending upon the urgency of the group's problems. This measure is *negatively* related to cohesiveness and member satisfaction in the less urgent situations, but *positively* related to cohesiveness in the more urgent situations. This difference is statistically significant. Last, the chairman's permissiveness is positively related to the cohesiveness of the group only in the more urgent situations. This difference in correlations is also statistically significant.

DISCUSSION

Over all groups in the present population and in several different situations, leadership sharing is associated with lessened group cohesion and lessened satisfaction with the conference. Although we have no direct evidence on this matter, these results would appear to suggest that the present groups have an expectation maintaining that the designated leader, the chairman, is to be the major behavioral leader in the group. Group members performing leadership functions are then reacted to negatively if their behavior is seen as challenging the position of the chairman as the group's major behavioral leader. This hypothesis does not deny that there may be other causes for negative reactions to the sharing of leadership functions; members perform these functions when they attempt to influence the group, and the group may not agree that the influence attempt is in the proper direction. However, if there is the tradition we have hypothesized, it will tend to act like any social norm and nonconformists will be rejected.

Other investigations undertaken by the Conference Research project lend some support to this *post hoc* hypothesis. One of these was an interview study of executives in 75 business and governmental organizations; the other was an experiment carried out at the University of Michigan. One of the conclusions from the interviews was stated as follows: "Most of the executives felt that what occurs in a conference is primarily the responsibility of the leader. When asked what they thought were the important factors in obtaining a successful conference, 86 per cent of their replies were in terms of

factors which ordinarily are controlled by the leader before the conference begins" (10, p. 16). Few executives considered motivational factors. "The paucity of comment on this score clearly reflects the executive's conception of conferences as serving only his administrative needs" (10, p. 16). We may readily infer from this that the present population of groups can be characterized as having customary leader control and leader dominance.

Results from the experiment mentioned above also suggest a fairly widespread tradition concerning the role of the conference chairman. If there is such a tradition, one that has the properties of a group norm, we might predict that violators of the tradition would tend to be rejected by the group. There is evidence on this point in the experiment conducted by Heyns (6).

Heynes set up two divergent leadership styles: "positive" and "negative" leadership. The "positive" leader performed functions related to group maintenance and goal attainment while the "negative" leader not only failed to perform these functions, but attempted "to create the impression that one part of the group was working at cross purposes with other parts." The obtained correlations suggest, in part, that when the designated leader does perform leading functions, performance of these behaviors by members other than the leader tends to result in these members being rated by the group less acceptable both as a person and as a participant. However, when the leader is inactive and nonhelpful, "responsibility for behaviors which are accepted as leader behaviors tends to make one perceive unity and feel accepted both as a person and as a participant" (6, p. 115). In other words, those members who acted like leaders when the designated leader performed his functions tended to be rejected, while they were not rejected when the designated leader failed to perform the functions associated with his role.

In the two populations considered, then, college students in experimental situations and executives in business, government, and industry, the groups behaved *as if* they wanted the designated leader to be the sole major behavioral leader. In somewhat different terms, they may have expected a role differentiation with the chairman's role being clearly differentiated from the members' roles. The conditions under which the field study and experiment were conducted suggest a possible limiting condition for this expectation. In both cases the designated leaders were clearly of a higher status level than the majority of the group members. (The departmental chairman was the leader of the experimental groups.) Desire to maintain this hierarchical structure—not to have rivalry with the high status leader—may thus account for the present results.

The evidence reported in this paper also suggests two types of conditions that may affect the group's reactions to violations of this expectancy. In the Heyns' experiment there are indications that the hypothesized group tradition becomes less important when the designated leader does not perform the functions associated with his role. The members who attempt to fill the "gap" left by the chairman are not rejected by the group.

This expectancy may also be less important when there are urgent problems confronting the group, according to the results of the field study presented above. It will be recalled that cohesiveness and satisfaction tend to decrease the less differentiated the chairman is from the other group members, the more the members participate relative to the chairman, and the less the chairman controls the group's process—but only when the group's problems are not too urgent. On the other hand, there are positive relationships with the permissiveness of the leader and the percentage of all the solution proposing done by the members when the group's problems are relatively urgent. The group's motivation to reach a problem solution as quickly as possible thus appears to be stronger than its motivation to conform to the expectancies concerning role differentiation. Not only does leadership sharing fail to lessen cohesiveness and member satisfaction in the more urgent conditions, but the leader's permissiveness and the proposing of solutions by the members tends to make for more attractive group situations.

The present results point up the necessity of considering the relational aspects of leadership. In the present case, the effects of leadership sharing cannot be fully understood without taking into consideration the

relationship between this sharing and the expectancies and motivations of the group.

SUMMARY

Seventy-two groups of small decision-making conferences in business, industry, and government were studied by a team of observers. The conference participants filled out a brief questionnaire upon the completion of the meeting and were later interviewed. Correlational analysis was applied to the data obtained from the observers and the conference participants. The following hypotheses were developed to explain the results:

1. There is a general expectation in the present population of groups maintaining that the socially designated leader, the chairman, should be the sole major behavioral leader. In other words, there appears to be an expectancy of role differentiation between the designated leader and the group members with each performing somewhat unique functions.

a. Leadership sharing by members other than the designated leader tends to be related to a decrease in group cohesiveness and satisfaction with the meeting over the entire sample of groups, and in groups with more and less permissive leaders.

b. These results also hold in groups contrasted on whether the leadership sharing is generally supporting or less supporting of the chairman.

2. Leadership sharing is not generally reacted to negatively in groups with urgent problems. It may be that the group's motivation to reach an adequate problem solution as quickly as possible lessens the importance of the hypothesized group tradition.

Some other investigations undertaken by the Conference Research project are cited to support the hypothesis of an expectancy of role differentiation in the present population.

REFERENCES

1. Conference Research project. Dittoed reports, University of Michigan, 1948–1951.
2. COWLEY, W. H. Traits of face-to-face leaders. *J. abnorm. soc. Psychol.*, 1931, **5**, 304–313.
3. FESTINGER, L., *et al. Theory and experiment in social communication.* Ann Arbor, Michigan: Univer. of Michigan Press, 1950.
4. FRENCH, J. R. P., JR. Organized and unorganized groups under fear and frustration. *On authority and frustration: Studies in topological and vector psychology.* III. Iowa City: State Univer. of Iowa, 1944.
5. FRENCH, R. L. Leadership and the authoritarian-democratic dimension. Paper read at Amer. Psychol. Ass., State College, Pa., September, 1950.
6. HEYNS, R. W. Effects of variation in leadership on participant behavior in discussion groups. Unpublished doctoral dissertation, Univer. of Michigan, 1948.
7. HEYNS, R. W. Functional analysis of group problem-solving. Dittoed report Conference Research, Univer. of Michigan, 1948.
8. JENNINGS, H. In T. M. Newcomb and E. L. Hartley (Eds.), *Readings in social psychology.* New York: 1947. Pp. 407–412.
9. KATZ, D. Social psychology and group processes. In C. P. Stone (Ed.), *Annual review of psychology.* Vol. II. Stanford: Annual Reviews, 1951. Pp. 137–142.
10. KRIESBERG, M. Executives evaluate administrative conferences. *Advanced Management,* March, 1950, 15–17.
11. MARQUIS, D. G., GUETZKOW, H., HEYNS, R. W. A social psychological study of the decision-making conference. In H. Guetzkow (Ed.), *Groups, leadership and men.* Pittsburgh: Carnegie Press, 1951. Pp. 55–67.
12. SANFORD, F. H. *Authoritarianism and leadership.* Philadelphia: Institute for Research in Human Relations, 1950.
13. SHARTLE, C. L. Leadership and executive performance. *Personnel,* March, 1949, 370–380.

Received May 7, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

CONCEPT OF "SUBTLETY" IN STRUCTURED PSYCHIATRIC AND PERSONALITY TESTS: AN EXPERIMENTAL APPROACH[1]

WILLIAM SEEMAN
Mayo Clinic

IT is the primary purpose of those tests, structured[2] or unstructured, which are designed with clinical utilitarian objectives in mind, to provide the clinician with such information as will make it possible for him to predict those facets of non-test behavior which he regards as interesting, relevant, or significant. That is, the clinician has little or no intrinsic interest in the number of movement responses of a Rorschach protocol, the number of suicides in TAT[3] stories, the magnitude of elevations of MMPI[4] scales, or the profile pattern of any other test instrument. Such information is of interest only to the extent that these aspects of test products can be shown to have some predictable relation to the non-test behavior or behavior dispositions which *are* of intrinsic interest to the clinician. That is, the information gleaned from these tests is of value only in so far as it contributes to the ability of the psychologist more accurately to describe the behavior tendencies of the individual. This is a simple and obvious notion which appears, however, to have been somewhat more clearly understood about projective (unstructured) tests than about structured, questionnaire-type instruments. The authors of early structured inventories appear to have considered the items of the test themselves of intrinsic importance, since these items seem to have been regarded as *prima facie* samples of the behavior in which the test maker was interested. Thus, as Meehl (3) commented, if the test makers wanted to know whether a man blushed frequently they took the position that "We cannot drop him into a social situation to observe whether he blushes, so we do the next best thing and simply ask him, 'Do you blush easily?'" Meehl continued:

> Associated with this approach to structured personality tests is the construction of items and their assembling into scales upon an *a priori* basis, requiring the assumption that the psychologist building the test has sufficient insight into the dynamics of verbal behavior and its relation to the inner core of personality that he is able to predict beforehand what certain sorts of people will say about themselves when asked certain sorts of questions (3, p. 297).

With the growing importance of projective tests it became fashionable to distinguish between these and structured instruments on the ground that it is a fundamental defining property of the structured instrument that "each of the test questions will have the same meaning to all subjects who take the examination" (2, p. 135), whereas the stimulus situation in the case of projective instruments is designed so that "it will mean to the subject, not what the experimenter has arbitrarily decided it should (as in most psychological experiments using standardized stimuli in order to be 'objective'), but rather whatever it must mean to the personality who gives it . . . his private, idiosyncratic meaning . . ." (1, p. 403). Recently, however, Meehl (3) had occasion, in a cogently reasoned paper, to question these several assumptions. He argued (*a*) that although most test makers built their structured instruments in terms of an a priori assignment of identical meaning to items, this was "not by any means a part of their essential nature"; (*b*) that it was not necessary to consider the items as "feeble surrogates for a behavior sample," and (*c*) that the items of a structured test might be regarded, rather, as stimuli "the non-test correlates of which must be discovered by empirical means."

This relatively more recent view of structured psychiatric and personality instruments has some striking consequences for the notion of *meaning* of items; consequences which it is the purpose of this paper to present in

[1] The assistance of Dr. Paul E. Meehl, of the University of Minnesota, in securing the raw data on which this paper is based, is hereby acknowledged.

[2] For the purposes of this paper the term *structured* is intended to convey the notion of a primarily *verbal* test.

[3] Thematic Apperception Test.

[4] Minnesota Multiphasic Personality Inventory.

experimental terms. For it is a view which opens the way for *subtlety* in such tests. This notion of subtlety may perhaps be best developed by quoting from a previous experimental paper by the writer:

> It is the "meaning" property of the stimuli in structured personality tests which is the subject of this investigation, and the term "subtlety" will be used to refer to the degree to which this "meaning" can or cannot be arbitrarily assigned in *a priori* fashion. As an example consider two items from the MMPI: "It takes a lot of argument to convince most people of the truth" and "I have a habit of counting things that are not important, such as bulbs on electric signs and so forth." To the extent that the *psychodynamic meaning* of the second item can easily be established with a high degree of interperson agreement (that is, most individuals who have had the requisite training would agree that this is an obsessive-compulsive mode of defense) whereas this is not true of the first item, the first item would *by definition* be properly characterized as "more subtle" than the second. For experimental purposes this definition will prove adequate although it does not provide for a complete ordering of a hierarchy of "subtlety." . . . It seems clear, however, that "subtlety" is a quality which cannot be dichotomized in all-or-none fashion, but rather that there are *degrees* of this property (5, p. 278).

In the paper, part of which has just been quoted, evidence was presented to show that 15 items, selected from the MMPI as subtle in character, behaved in a strikingly different manner from 15 items characterized as obvious. In that study, 58 students in clinical psychology who had previously taken two quarters of abnormal psychology among other prerequisite courses, were asked to assign the specified items to a psychiatric scale or to indicate the direction (true, false, or cannot say) in which certain types of patients would respond to the items. The evidence of that study was regarded as giving experimental meaning to the concept of subtlety and unequivocally supporting the concept.

The present paper provides a further experimental test of the concept which may, in a sense, be characterized as the acid test, for it presents results of the readministration of the same items to the same group of students after completion of an additional two quarters of clinical psychology at the University of Minnesota. Among the circumstances which render these results of special significance may be mentioned the fact that almost an entire quarter is devoted to a study of the MMPI, the instrument from which all these items are culled, as well as the fact that these lectures are given by Dr. Paul Meehl, whose familiarity with the test probably approximates that of the test's author.

METHOD

All items subjected to experimental investigation were, as has been indicated, taken from the MMPI. There were 30 such items, 15 of which were designated as subtle because it was considered that their "meaning" and psychological significance would not be detected, even after one quarter of intensive study of the MMPI. The other 15 items were designated as obvious items; these were items such as the "counting" item previously referred to, whose "psychological meaning" was considered to be quite clearly recognizable by psychologically trained individuals. The 30 items, the manner of presentation, and the directions are described below.

Part 1

Directions. This is an experiment in which your assistance is being enlisted. The problem is to identify items which, in your opinion, would be of diagnostic value on some psychiatric scale of a structured type of psychological test. A "psychiatric scale" is to be regarded as a category such as "schizophrenia," "hysteria," "depression," and so forth.

Read each item and decide whether the answer "true" or "false" to that item would render the item of value in any of the psychiatric categories listed with it. If you decide it belongs on a certain diagnostic scale, put an X before that category. If you think the item would be of no value for any of the categories listed, place the X before "none." As an example consider the item: "I am happy most of the time." The answer "false" would help to identify this as a depressed person and so would belong on the "Depression" scale of the test.

The items in part 1 are as follows:

1. I believe I am being plotted against. ——a. Depression ——b. Obsessive-compulsive ——c. Paranoid schizophrenia ——d. Hysteria ——e. None
2. I have had periods in which I carried on activities without knowing later what I had been doing. ——a. Schizophrenia ——b. Paranoid state or condition ——c. Hysteria ——d. Obsessive-compulsive ——e. None
3. Peculiar odors come to me at times. ——a. Hypomania ——b. Schizophrenia ——c. Depression ——d. Obsessive-compulsive ——e. None
4. Much of the time my head seems to hurt all over. ——a. Paranoid state ——b. Obsessive-compulsive ——c. Depression ——d. Hysteria ——e. None
5. I am neither gaining nor losing weight. ——a. Hysteria ——b. Depression ——c. Hypomania ——d. Schizophrenia ——e. None
6. It takes a lot of argument to convince some people of the truth. ——a. Paranoid schizophrenia ——b.

Obsessive-compulsive ——c. Paranoid state ——d. Hysteria ——e. None

7. Most people will use somewhat unfair means to gain profit or an advantage rather than lose it. ——a. Hysteria ——b. Depression ——c. Hypomania ——d. Schizophrenia ——e. None
8. Most nights I go to sleep without thoughts or ideas bothering me. ——a. Hypomania ——b. Paranoid schizophrenia ——c. Obsessive-compulsive ——d. Hysteria ——e. None
9. I almost never dream. ——a. Depression ——b. Obsessive-compulsive ——c. Paranoid state ——d. Hypomania ——e. None
10. I often feel as if things were not real. ——a. Hysteria ——b. Schizophrenia ——c. Hypomania ——d. Paranoid state ——e. None
11. I have a habit of counting things that are not important such as bulbs on electric signs and so forth. ——a. Obsessive-compulsive ——b. Hysteria ——c. Schizophrenia ——d. Mania ——e. None
12. Someone is hypnotizing me and has control of my mind. ——a. Depression ——b. Hysteria ——c. Obsessive-compulsive ——d. Paranoid schizophrenia ——e. None
13. I commonly wonder what hidden reason a person may have for doing something nice for me. ——a. Obsessive-compulsive ——b. Hypomania ——c. Hysteria ——d. Depression ——e. None
14. Sometimes some unimportant thought will run through my mind and bother me for days. ——a. Obsessive-compulsive ——b. Hypomania ——c. Schizophrenia ——d. Depression ——e. None
15. I get mad easily and then get over it. ——a. Hysteria ——b. Paranoid state ——c. Hypomania ——d. Obsessive-compulsive ——e. None
16. I am very seldom troubled by constipation. ——a. Hysteria ——b. Depression ——c. Hypomania ——d. Obsessive-compulsive ——e. None
17. I like to flirt. ——a. Hysteria ——b. Paranoid schizophrenia ——c. Depression ——d. Hypomania ——e. None
18. I can be friendly with people who do things which I consider wrong. ——a. Obsessive-compulsive ——b. Depression ——c. Hypomania ——d. Hysteria ——e. None

Part 2

Directions. In each of the items below indicate by encircling the appropriate character how the item should be scored for the category mentioned. In every instance T means *True*, F means *False*, and ? means *Cannot Say*. For example, the item "I seem to be as capable and smart as most others around me," would be scored as a depressed answer if answered F.

1. The item "most people inwardly dislike putting themselves out to help other people" would be scored as a paranoid answer if answered T F ?
2. The item "I wish I could be as happy as others seem to be" would be scored as a depressed answer if answered T F ?
3. The item "I never worry about my looks" would be scored for hypomania if answered T F ?
4. "I dream frequently about things that are best kept to myself" would be scored as depressive if answered T F ?
5. "Someone has been trying to influence my mind" would be scored for paranoid schizophrenia if answered T F ?
6. "I do not blame a person for taking advantage of someone who lays himself open to it" would be scored in the direction of depression if answered T F ?
7. "Some people are so bossy that I feel like doing the opposite of what they request even though I know they are right." This item would be scored as paranoid if answered T F ?
8. "Much of the time I feel as if I have done something wrong or evil" would be scored as obsessive-compulsive if answered T F ?
9. "I usually have to stop and think before acting even in trifling matters" would be scored as obsessive-compulsive if answered T F ?
10. "I drink an unusually large amount of water every day" would be scored for hysteria if answered T F ?
11. "Bad words, often terrible words, come into my mind and I cannot get rid of them." This item would be scored as obsessive-compulsive if answered T F ?
12. "Most people are honest chiefly through fear of being caught" would be scored in the paranoid direction if answered T F ?

Results

The primary question to be answered by the results of this experimental treatment is whether, after a quarter of study of the MMPI, the items designated as subtle maintain their subtlety. This question will be answered, of course, by comparing the behavior of the two types or sets of items. The failure of the subtle items to behave in a discriminably different manner from the obvious items would indicate that whatever the degree of subtlety inherent in the former set of items, it was not sufficient to withstand the instruction in the course work. On the other hand, the concept of subtlety may be said to receive experimental verification if such differences *can* be demonstrated. The definition of subtlety as developed here requires, rather obviously, that a significantly greater success be achieved by the subjects in the categorizing of the items as well as in indicating the direction of scoring for the specified category. One might expect, also, that the "none" category be used with a significantly greater frequency for the subtle than for the obvious items, either on the ground that the former are innocuous or because of the ambiguity of their "meaning." One might expect, also, that occasionally a *chance* assignment of the subtle items should occur, in consequence of the ambiguity of their "meaning." Finally, if one were to consider the 15 subtle and the 15 obvious items as two separate subtests, one would expect a significantly different distribution of scores, considering the number of correct identifications as a score.

Table 1 indicates the manner in which the items in Part 1 (the first 18 items) were assigned to the various psychiatric categories. The MMPI criterion scale is, in each case, indicated by an asterisk, so that the number in that cell indicates the number of correct identifications of the item; that is, the number of students who assign the item to the same scale as the criterion. It is quite evident from a study of the table that the degree of success achieved in assignment of the obvious items is markedly greater than that achieved in assignment of the subtle items.

TABLE 1

SUCCESS AND FAILURE IN IDENTIFYING APPROPRIATE PSYCHIATRIC SCALES AS ACHEIVED BY 48 CLINICAL PSYCHOLOGY STUDENTS AFTER COURSE DISCUSSION ON MMPI †

ITEM **	*D*	*Hy*	*Pa*	*O–C*	*PaSc*	*Sc*	*Ma*	NONE	*p* ‡
1					47*			1	.01
2		23	1	2		21*		1	.01
3	1			4		39*		4	.01
4	8	29*		1				10	.01
5(S)	7	19*					3	18	.01
6(S)		3	21	22	2				.01
7(S)	7	20*					3	18	.01
8		3		34*	1		2	8	.01
9(S)	2			13*	10		4	19	.01
10		1				47*			.01
11				48*					.01
12				1	47*				.01
13(S)	10	6		13				19	.01
14	1			45*		1		1	.01
15(S)		18*		2			24	4	.01
16(S)	13*	19*		4			4	8	.01
17(S)	4*	10			3		25	6	.01
18(S)	1	22*		7			11	7	.01

† Depression, hysteria, paranoid state or condition, obsessive-compulsive, paranoid schizophrenia, schizophrenia, mania or hypomania.

‡ Probability of obtaining the indicated distribution.

* Indicates appropriate scale using MMPI as criterion.

** (S) indicates "subtle" items.

This is in agreement with the concept of subtlety previously developed, and suggests that even instruction in the MMPI by an individual thoroughly conversant with its structure does not close the gap between the subtle and the obvious items. If one arranges the responses in a contingency table indicating the agreement of choices with criterion scales and computes mean square contingency coefficients, one finds that for the obvious items $C=.83$, whereas for the subtle items, $C=.50$; a striking difference.[5] Or, alternately, if one selects an arbitrary criterion of 75 per cent correct for each item, the resulting differences (Table 2) between the two sets of items are significant at the .01 level. Furthermore, the difference between the two sets of items for the "none" category is 17.2 per cent, a difference significant at the .011 level. Although this is a significant difference and in the required direction, it can certainly not be claimed as a dramatic difference; however, there are, as will be seen subsequently, a number of quite dramatic differences.

[5] It is recognized that these contingency coefficients involve judgments which were not of an independent character. Consequently, alone, they would not be acceptable as evidence. However, they are buttressed by other evidence quite unequivocal in nature.

TABLE 2

BREAKDOWN OF SUBTLE AND OBVIOUS ITEMS IN PARTS 1 AND 2 ACCORDING TO WHETHER 75 PER CENT SUCCESS WAS ACHIEVED OR NOT

DEGREE OF SUCCESS	SUBTLE	OBVIOUS
>75%	0	12
<75%	15	3
	$\chi^2=20$	
	$p<.01$	

Two such differences are presented in Figs. 1 and 2, wherein are charted the item-by-item successes achieved on both types of

items in Part 1 and Part 2. It is clear from Fig. 1 that only in the case of a single subtle item (No. 18) does the success in assignment of items to categories exceed the success achieved on the *lowest* of the obvious items (No. 2). Even more striking is the difference between the items as graphed in Fig. 2, for here it is observed that the *greatest* success achieved in a subtle item (No. 6) is well below the most feeble success on the subtle items (No. 3). Furthermore, analysis of the cumulative success for all the items (Fig. 3) reveals again some striking differences, all of the kind which are consistent with and indeed required by, the concept of subtlety developed in this paper.

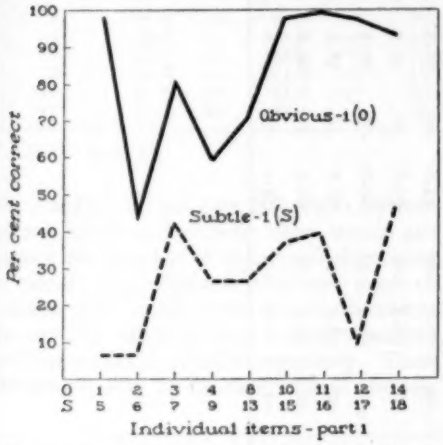

FIG. 1. PER CENT SUCCESS ACHIEVED ON EACH OF THE "OBVIOUS" AND "SUBTLE" ITEMS IN PART 1

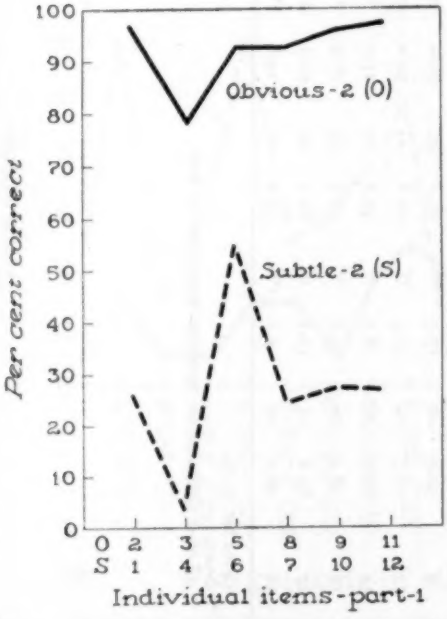

FIG. 2. PER CENT SUCCESS ACHIEVED ON EACH OF THE "OBVIOUS" AND "SUBTLE" ITEMS IN PART 2

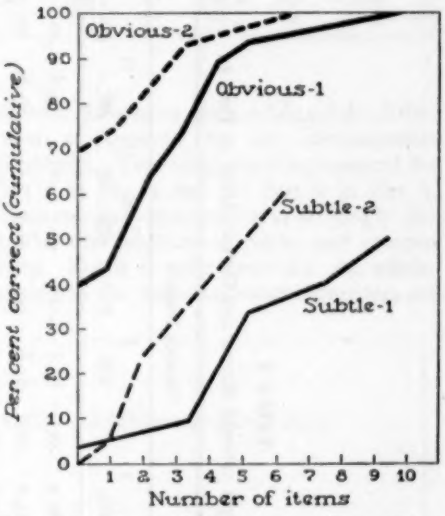

FIG. 3. CUMULATIVE SUCCESS FOR ALL "SUBTLE" AND ALL "OBVIOUS" ITEMS

Still another way of demonstrating the considerable difference in behavior between subtle and obvious items is illustrated in Tables 3 and 4. Here are presented all differences for all possible combinations of subtle versus obvious items; that is, Table 3 presents differences between all possible subtle-obvious pairs in Part 1 and Table 4, differences between all possible subtle-obvious pairs in Part 2. This yields 9×9 or 81 subtle-obvious differences for Part 1 and 6×6 or 36 such differences for Part 2. Now we may reason that if the obvious and the subtle items were not, in fact, drawn from different populations (that is, if they were not different *kinds* of items), the number of paired differences which are actually significant at the .01 level should not exceed the number which would be obtained at this level by chance. Or, perhaps, it would be wiser to put it that

TABLE 3

DISTRIBUTION OF DIFFERENCES BETWEEN SUBTLE AND OBVIOUS ITEMS IN PART I

SUBTLE ITEMS	OBVIOUS ITEMS																	
	1		2		3		4		8		10		11		12		14	
	diff.	*p*	*diff.*	*p*	*diff.*	*p*	*diff.*	*p*	*diff.*	*p*	*diff.*	*p*	*diff.*	*p*	*diff.*	*p*	*diff.*	*p*
5	91.7	.01	37.5	.01	75.0	.01	54.2	.01	64.6	.01	91.7	.01	93.8	.01	91.7	.01	87.5	.01
6	91.7	.01	37.5	.01	75.0	.01	54.2	.01	64.6	.01	91.7	.01	93.8	.01	91.7	.01	87.5	.01
7	56.3	.01	2.1	**	39.6	.01	18.8	**	29.2	.01	56.3	.01	58.4	.01	56.3	.01	52.1	.01
9	71.5	.01	17.3	**	54.8	.01	34.0	.01	44.4	.01	71.5	.01	73.6	.01	71.5	.01	67.3	.01
13	71.5	.01	17.3	**	54.8	.01	34.0	.01	44.4	.01	71.5	.01	73.6	.01	71.5	.01	67.3	.01
15	60.5	.01	6.3	**	43.6	.01	23.0	*	33.4	.01	60.5	.01	62.6	.01	60.5	.01	56.3	.01
16	58.4	.01	4.2	**	41.7	.01	20.9	*	31.3	.01	58.4	.01	60.5	.01	58.4	.01	54.2	.01
17	89.6	.01	35.4	.01	72.9	.01	52.1	.01	62.5	.01	89.6	.01	91.7	.01	89.6	.01	85.4	.01
18	52.1	.01	2.1	**	35.4	.01	14.6	**	25.0	.01	52.1	.01	54.2	.01	52.1	.01	47.9	.01

* $>.01$ but $<.05$.
** $>.05$.
All other differences are significant beyond the .01 level.

TABLE 4

Distribution of Differences between Subtle and Obvious Items in Part 2

Subtle Items	Obvious Items											
	2		3		5		8		9		11	
	diff.	*p*	*diff.*	*p*	*diff.*	*p*	*diff.*	*p*	*diff.*	*p*	*diff.*	*p*
1	71.5	.01	52.7	.01	67.3	.01	67.3	.01	69.4	.01	71.5	.01
4	95.8	.01	77.0	.01	91.6	.01	91.6	.01	93.7	.01	71.5	.01
6	35.4	.01	22.9	*	37.5	.01	37.5	.01	39.6	.01	95.8	.01
7	72.9	.01	54.1	.01	68.7	.01	68.7	.01	70.8	.01	35.4	.01
10	71.5	.01	52.7	.01	67.3	.01	67.3	.01	69.4	.01	72.9	.01
12	71.5	.01	52.7	.01	67.3	.01	67.3	.01	69.4	.01	71.5	.01

* $>.01$ but $<.05$.

the number of 1 per cent differences between paired subtle and obvious items should not exceed the number of differences significant at the .01 level obtained when one takes all possible pairs of *like* items, that is, differences among the obvious items taken separately and the subtle items taken separately. These differences may be characterized as obvious-obvious differences and subtle-subtle differences as opposed to the obvious-subtle differences. This comparison is presented for Part 1 in Fig. 4 and for Part 2 in Fig. 5. Again we may characterize as strikingly dissimilar the behavior of subtle and obvious items. For it is quite clear that the subtle-subtle and the obvious-obvious differences are

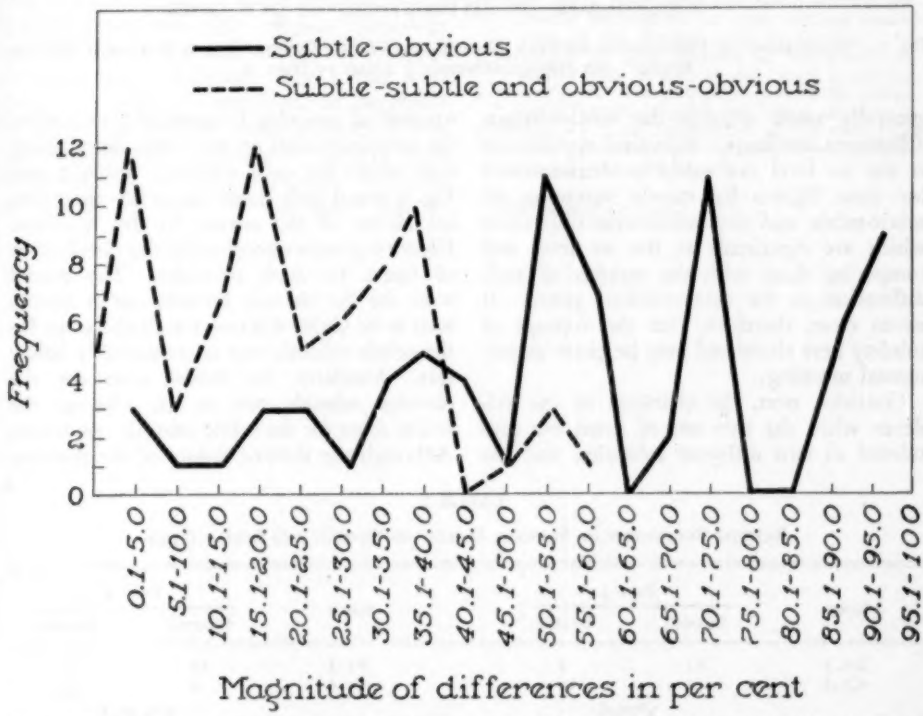

Fig. 4. Comparison of Differences between Dissimilar ("Subtle-Obvious") and Similar ("Subtle-Subtle" and "Obvious-Obvious") Items in Part 1

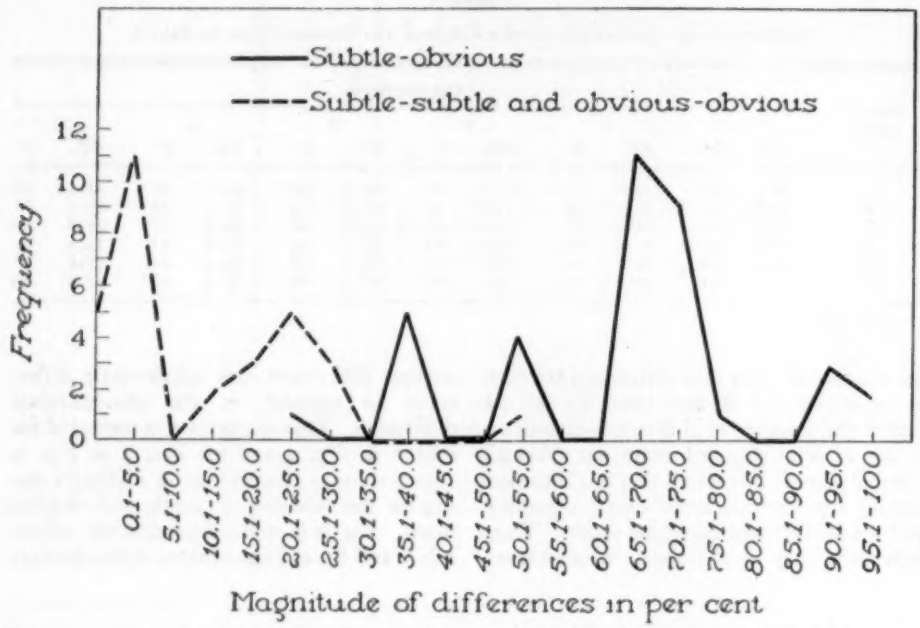

FIG. 5. COMPARISON OF DIFFERENCES BETWEEN DISSIMILAR ("SUBTLE-OBVIOUS") AND SIMILAR ("SUBTLE-SUBTLE" AND "OBVIOUS-OBVIOUS") ITEMS IN PART 2

generally small whereas the subtle-obvious differences are large. Statistical significance at the .01 level can easily be demonstrated for these figures by merely summing all subtle-subtle and obvious-obvious differences which are significant at the .01 level and comparing these with the number of such differences in the subtle-obvious group. It seems clear, therefore, that the concept of subtlety here elucidated may be given experimental meaning.

Consider, next, the character of the evidence when the two sets of items are considered as two different subscales, and the number of successes is considered as a score for each individual on each subscale, yielding four scores for each student. Table 5 and Fig. 6 reveal little doubt about the unequivocal nature of the answer to this question. Figure 6 presents graphically the distribution of scores for both subscales. The modal score for the obvious subscale one is readily seen to be eight, whereas the modal score for the subtle subscale one is considerably below this. Similarly, the modal score for the obvious subscale two is six, whereas the modal score for the subtle subscale two is one. Although the skewed nature of the distribu-

TABLE 5

RANDOM BREAKDOWN OF SUBSCALE SCORES FOR OBVIOUS AND SUBTLE ITEMS

SCORE	PART 1		SCORE	PART 2	
	OBVIOUS	SUBTLE		OBVIOUS	SUBTLE
>5.5	23	0	>3.5	22	1
<5.5	1	24	<3.5	2	23
	$\chi^2=46$			$\chi^2=36.8$	
	$p<.01$			$p<.01$	

tion renders it incorrect to use normal curve procedures, it can, nevertheless, readily be shown by use of nonparametric statistics that these distributions are different, significant at the .01 level. The procedure followed in this

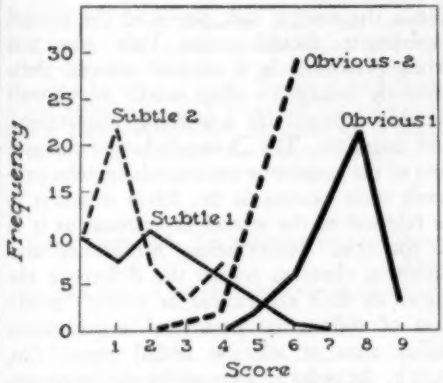

FIG. 6. DISTRIBUTION OF SCORES ON "OBVIOUS" AND "SUBTLE" SUBSCALES

case was the following: The entire group was broken down into eight random subgroups and the means for the subtle and obvious subscales were computed. These were then ranged in order of magnitude as suggested by Moses (4), and the Mann-Whitney *U* test was applied, the resulting *U* falling well beyond the .01 level of significance. It is possible, also, to apply the chi-square test of significance provided precautions are taken to insure independence of categories. This was done by assigning the members of the entire sample in random fashion to two subgroups. The chi square was then computed with the criteria indicated in Table 5, the difference being in the required direction and well beyond the .01 level of significance.

It seems clear from all the evidence reviewed above that the concept of subtlety in structured tests can be given solid experimental meaning. It has been demonstrated that those items designated as subtle not only behaved in a discriminably different fashion on the occasion of their initial presentation, but that this difference was manifested after the students were given additional instruction, virtually one full quarter, which consisted of lectures on the structure of the MMPI by an unusually competent lecturer.

REFERENCES

1. FRANK, L. K. Projective methods for the study of personality. *J. Psychol.*, 1939, **8**, 389–413.
2. HUTT, M. L. The use of projective methods of personality measurement in Army medical installations. *J. clin. Psychol.*, 1945, **1**, 134–140.
3. MEEHL, P. E. The dynamics of "structured" personality tests. *J. clin. Psychol.*, 1945, **1**, 296–303.
4. MOSES, L. E. Non-parametric statistics for psychological research. *Psychol. Bull.*, 1952, **49**, 122–143.
5. SEEMAN, W. An investigation of the concept of "subtlety" in structured personality tests. *J. consult. Psychol.*, 1952, **16**, 278–283.

Received May 12, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

MANIFEST ANXIETY AND "PERCEPTUAL DEFENSE"

M. E. BITTERMAN AND CALVIN W. KNIFFIN
University of Texas

In a recent issue of this JOURNAL, McGinnies and Sherman (10) described an experiment purported to provide a clear demonstration of "perceptual defense." Tachistoscopic thresholds for eight affectively neutral "task words" were determined. Four of these words were always exposed following 2-sec. presentations of "taboo" words, while exposures of the other four always followed presentations of neutral words. The significantly higher thresholds for the task words which followed taboo words were attributed by the authors to the "generalization of perceptual defense."

Now it is true that this experimental design is not open to one important criticism which Howes and Solomon (5) made of an earlier attempt by McGinnies (8) to demonstrate the reality of "perceptual defense"—the procedure employed seems to rule out the possibility that the obtained difference in thresholds can be understood in terms of the frequency of previous experience with the task words. It is by no means clear, however, that the new design avoids another major criticism suggested by Howes and Solomon—the higher thresholds for task words following taboo words may be attributed to hesitancy in reporting rather than to an immediate effect on the perceptual process. There is no question but that reluctance to report plays an important part in direct determinations of thresholds for taboo words; "the foul plot to deceive the experimenters" to which McGinnies has made skeptical reference (9, p. 240) could have been discovered in Tuscaloosa as well as in Cambridge by the simple expediency of questioning the observers.[1]

[1] Excerpts from the reports of *Ss* asked in general terms to comment on their experiences during such an experiment: "I hesitated to report the first one because I didn't expect to be shown that type of word." "I saw the first one, I know, a few times before I reported it." "I wanted to be sure I saw every letter of those words before I called them." "I saw *kotex* but kept telling myself it can't be that." "The first word was a surprise, and I hesitated to call what it looked like." "I guessed at the other words but on those I didn't report until I saw every letter in them." "I definitely saw the words but thought 'Oh no, it couldn't be that' and kept trying to see something else." "I identified *raped* at the very first, but hesitated to say it."

While the neutral task words of the second experiment should create little suspicion when presented in a neutral context, their unsubtle linkage to taboo words might well provide grounds for hesitation, uncertainty, and restraint. The Howes-Solomon description of the suspicious undergraduate who proceeds with caution in the taboo situation is as relevant to the second experiment as it is to the first. Nevertheless, McGinnies and Sherman chose to regard the difference obtained in their experiment as a clear "reflection of differential perceptual organization rather than of selective verbal report" (10, p. 81). In order to account for the "generalization of defense" they assumed that the taboo words "signal a state of approaching punishment and consequently become cues for eliciting the anxiety associated with actual punishment" (10, p. 84). The anxiety engenders avoidance reactions which result in elevated recognition thresholds and which may persist long enough to interfere with the perception of subsequent neutral words that do not in themselves initiate defensive behavior.

Even if we assume that hesitancy in reporting played some role in these experiments on "perceptual defense," it is conceivable that the differences obtained may have been due *in part* to perceptual blocking. While it is difficult to isolate the two variables experimentally, evidence is available which bears on the reinforcement theory of the hypothetical defensive process proposed by McGinnies and Sherman. In the course of a systematic investigation of the correlates of Taylor's Manifest Anxiety Scale (13, 14) under way in this laboratory, there was occasion to study the performance of criterion groups selected on the basis of that scale in a situation corresponding to the one employed in McGinnies' first experiment. The data obtained have direct implications for the validity of the anxiety interpretation proposed by McGinnies and Sherman as well as indirect implications for the question of whether it is necessary to postulate a defensive perceptual process in

order to account for the results of these experiments.

METHOD

Subjects

From a group of 348 undergraduate women at the University of Texas to whom the complete MMPI had been administered, 40 subjects (*Ss*) were selected on the basis of scores on Taylor's scale of Manifest Anxiety (13, 14).[2] Twenty *Ss* with Taylor scores between 1 and 8 comprised the low-anxiety group, while the remaining *Ss*, whose scores ranged from 26 to 44, comprised the high-anxiety group.

RESULTS AND DISCUSSION

The relation between thresholds and anxiety level. Mean thresholds for the four taboo words and the four neutral controls, individually and combined, are shown in Table 1. For both criterion groups the difference between the combined thresholds for taboo and neutral words was found to be significant beyond the .01 level of confidence, results which are in accord with those obtained in the first experiment of McGinnies.

TABLE 1

MEAN THRESHOLDS (IN HUNDREDTHS OF A SECOND) FOR TABOO AND NEUTRAL WORDS IN RELATION TO ANXIETY LEVEL

GROUP	*a'*	*b'*	*c'*	*d'*	*a*	*b*	*c*	*d*	COMBINED TABOO	COMBINED NEUTRAL	*Diff.*
High-anxiety	8.6	7.0	6.6	9.2	4.4	4.3	4.7	4.0	7.8	4.3	3.5*
Low-anxiety	10.2	6.5	8.9	8.0	5.5	4.4	4.8	5.0	8.4	4.9	3.5*
*Diff.***	−1.6	0.5	−2.3	1.2	−1.1	−0.1	−0.1	−1.0	−0.6	−0.6	0.0

* Significant beyond the .01 level of confidence by Wilcoxon's nonparametric test for paired replicates (15).
** Neither the *t* test nor Wilcoxon's test for unpaired replicates (15) provided any basis for rejecting the null hypothesis with reference to differences between groups.

Procedure

With a Gerbrands mirror tachistoscope, the recognition threshold of each of the following words (typewritten in elite caps) was determined:

PRACTICE	TABOO	NEUTRAL
p, apple	*a'*, kotex	*a*, clear
r, dance	*b'*, whore	*b*, music
	c', belly	*c*, glass
	d', raped	*d*, river

The *S* began with the two neutral practice words, *p* and *r*. For half the *Ss* in each criterion group, threshold measurements of the eight experimental words were made in the order *a'*, *a*, *b'*, *b*, *c*, *c'*, *d*, *d'*, while for the other half of the *Ss* in each group these determinations were made in the reverse order (*d'*, *d*, *c'*, *c*, *b*, *b'*, *a*, *a'*). The *Ss* were instructed simply that words would be exposed for gradually increasing periods of time and that they were to report what they saw on each trial. Determinations were made in the ascending order, beginning with an exposure time of .01 sec., which was increased by .01 sec. for each subsequent presentation until the word was correctly reported. A ready-signal was given before each exposure.

[2] The numbers of the items in the scale follow: 7, 13, 14, 18, 31, 32, 43, 67, 86, 107, 125, 142, 158, 163, 186, 190, 191, 217, 230, 238, 241, 242, 263, 264, 287, 288, 301, 317, 321, 322, 335, 337, 340, 352, 361, 371, 397, 407, 418, 424, 431, 439, 499, 506, 523, 528, 530, 549, 553. See also (14).

At no point, however, did a significant difference between the high- and low-anxiety groups appear. Actually, the mean thresholds of the low-anxiety group were slightly greater than those of the high-anxiety group both for the combined taboo and for the combined neutral words, although the differences failed to approach statistical significance. The results show no indication, therefore, of the relation between anxiety level and threshold which the McGinnies-Sherman interpretation suggests. The fault might be said to lie with the criterion of anxiety which was used, but there is now considerable evidence to suggest its validity. Scores on the Taylor scale, the items for which were selected on the basis of clinical judgment, correlate highly (4) with scores on an empirical scale recently constructed by Winne (16). Relations predictable from the reinforcement theory for which McGinnies and Sherman have expressed a preference also have been found between Taylor scores and simple conditioning (12, 13), discrimination (3, 11), and verbal learning in the face of repeated failure (6). The conclusion seems warranted, therefore, that anxiety plays no important role in the experi-

mental situation under scrutiny. It should be noted, of course, that a positive relation between anxiety and behavior in this experiment would not rule out the possibility that the threshold differences were determined by reluctance to report rather than by perceptual distortion. The lack of such a relation, however, suggests that the taboo words were not threatening enough to *S*s to warrant the assumption of "perceptual defense."

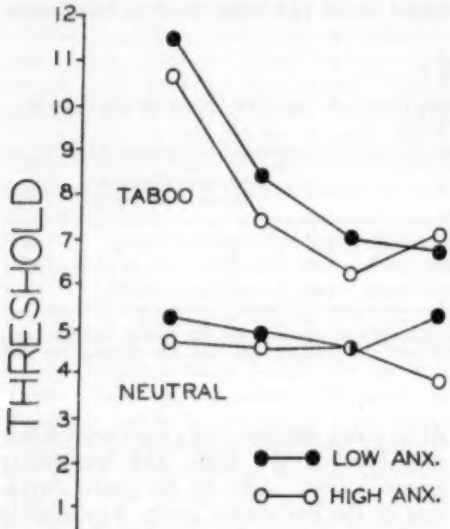

FIG. 1. MEAN THRESHOLDS (IN HUNDREDTHS OF A SECOND) FOR TABOO AND NEUTRAL WORDS AS A FUNCTION OF ORDER OF PRESENTATION

The order-effect. Because the postexperimental comments of *S*s suggested that reluctance to report was especially marked in the case of the first taboo word in the series, the data were analyzed for an order-effect. In Fig. 1, mean thresholds for taboo and neutral words are plotted as a function of the order in which the words of each category were presented. (Since order was counterbalanced in the experimental design, comparisons can be made between positions 1 and 4 and between positions 2 and 3 which will not be influenced by structural or experiential differences between words.) The curves for the two criterion groups are quite similar, and at no point do the differences between them approach statistical significance (t test and Wilcoxon's test for unpaired replicates). Although no order-effect is evidenced in the plots for neutral words, both groups show a marked decline in the thresholds for taboo words as a function of position in the series. For each group individually and for the two groups combined, the differences between positions 1 and 4 were significant beyond the .01 level of confidence (Wilcoxon's test for paired replicates [15]), although the differences between positions 2 and 3 fell somewhat short of significance in each case. The fact that no order-effect is evident for the neutral words suggests that the order-effect for the taboo words cannot be attributed to tachistoscopic practice. It must represent, instead, the development of insight into the essential nature of the situation. In the words of one *S*, "I caught on to what you were doing and then expected anything." This result is precisely what might be anticipated on the assumption that differences in the thresholds of taboo and neutral words (beyond those attributable to frequency of experience) may be accounted for in terms of differential readiness to report rather than in terms of perceptual functioning.

A preliminary analysis of the residual difference. Although the shape of the curve for the taboo words suggests that the order-effect rapidly disappears, there remains a difference between taboo and neutral words in position 4 which is significant beyond the .01 level of confidence (Wilcoxon's test for paired replicates [15]). How can this residual difference be accounted for? It is likely, as Howes and Solomon suggested, that frequency of experience with the words is an important variable, but the weight that should be assigned to this factor cannot be determined in the context of the present experiment. What is required for this purpose is an experiment like that of McGinnies and Sherman—any residual which which remains after repeated testing with comparable task words in such an experiment cannot be attributed to frequency of experience. The possibility also should be recognized that differential readiness to report continued to have some effect at position 4, and the role of perceptual distortion must be

considered. The importance of these factors can be estimated in a preliminary fashion from the data of the present experiment.

For every *S* the difference in thresholds for taboo and neutral words in position 4 was computed (threshold for fourth taboo word *minus* threshold for fourth neutral word), and on the basis of these scores (irrespective of anxiety scores) two groups of *Ss* were selected for consideration. The nine *Ss* who were least affected by the taboo word (whose difference scores were zero or negative) constituted the Low Group, and the eleven most affected *Ss* (whose difference scores ranged from 0.04 to 0.19 sec.) comprised the High Group.[8] For these *Ss* four additional scores, *K, Hy, Sc* (*K*-corrected), and *Pd* (*K*-corrected), were derived from the MMPI protocols from which the Taylor scores employed for the original selection of *Ss* were obtained.

The hypotheses underlying the choice of these indices were the following: The *K* and *Hy* scales were used to test the concept of "perceptual defense" on the assumption that it may be related to conversion disorders. The *Hy* scale measures the resemblance of responses on the test to those of hysterical patients (2). The *K* scale, which was designed to reflect a defensive, self-deceptive test-taking attitude, is highly correlated (.78 for normal females) with the subtle *Hy* items (7). The *Sc* scale was selected to test the hypothesis that perceptual defense is related to schizophrenic autism. A positive relation between taboo-neutral threshold difference and any of these scales would provide some support for the explanation of the threshold difference in terms of perceptual distortion. The Pd scale was selected as a measure of nonconformity to test the hypothesis that the residual difference is a function of readiness to report. On the assumption that the more unconventional subjects are less reluctant to report the taboo words, it was predicted that the Low Group would have a higher mean *Pd* score.

The results of this analysis are presented in Table 2. No significant differences were found for *K, Hy,* or *Sc* scores. Correlations between each of these scores and the fourth-position difference score for the entire group of 40 *Ss* (—.16, .19, and .21, respectively) also failed to attain significant proportions. These results, therefore, provide no support for the concept of "perceptual defense," although the validity of the psychometric criteria may be open to question. In the case of *Pd*, a significant difference did appear between the High and Low Groups, and there was a significant positive correlation (.44) between *Pd* scores and fourth-position difference scores for the entire group of 40 *Ss*. The relationship found was not, however, in the predicted direction; the higher the *Pd* score, the greater was the tendency for *S* to be affected by the taboo words. There seems to be no way of reconciling this finding with the concept of "perceptual defense," but a fairly reasonable interpretation (admittedly *ad hoc*) in terms of differential readiness to report can be found. The *Ss* were instructed to report whatever they thought they saw on each exposure, and conformity in the situation can be understood as obedience to instructions. It might be

TABLE 2

MEAN *K, Hy, Sc,* AND *Pd* SCORES IN RELATION TO THRESHOLD DIFFERENCES BETWEEN TABOO AND NEUTRAL WORDS

MMPI SCALE	FOURTH TABOO—FOURTH NEUTRAL			FIRST TABOO—FIRST NEUTRAL		
	LOW (*N*=9)	HIGH (*N*=11)	*Diff.*	LOW (*N*=12)	HIGH (*N*=9)	*Diff.*
K	15.3	13.5	−1.8	15.5	14.2	−1.3
Hy	22.4	24.7	2.3	22.5	25.0	2.5
Sc	25.0	31.6	6.6	25.7	30.1	4.4
Pd	19.5	25.1	5.6*	21.8	23.2	1.4

* Significant beyond the .01 level of confidence by Festinger's nonparametric method (1).

[8] The two groups did not differ with respect to anxiety level. Four of the nine *Ss* in the Low Group and six of the 11 *Ss* in the High Group belonged to the high-anxiety sample.

assumed, therefore, that, while all *Ss* were equally affected by the taboo words, the relatively independent, suspicious, and less pliable *Ss* (high *Pd* scores) did not feel constrained to report immediately and allowed themselves the privilege of another look at the situation. Whatever the validity of this interpretation, the significant relation between *Pd* score and threshold difference suggests that the residual difference cannot be accounted for entirely in terms of frequency of experience.

An analysis similar to that made for the fourth-position difference score was made for the first position, and the results are presented in Table 2. None of the four variables was found to be significantly related to the taboo-neutral difference at the first position. From these results it may be concluded that the initial effect was great enough to mask the personality differences which only later began to have a measurable influence. (All *Ss* show an initial reluctance to report the taboo words which is overcome more completely by those with low *Pd* scores.) A third analysis in terms of the difference between the mean threshold of all taboo words and the mean of all neutral words again showed the High Group to have a significantly higher mean *Pd* score (beyond the .05 level of confidence by Festinger's test [1]). All of these results should, however, be taken as extremely tentative, since they were derived secondarily from an experiment designed to study the role of anxiety. The primary results show clearly that the McGinnies effect is unrelated to anxiety as measured by the Taylor scale. The secondary results (on the order-effect, other psychometric correlates, and the comments of *Ss*) *suggest* that behavior in the experimental situation can be understood without the assumption of perceptual distortion.

SUMMARY

From a large sample of undergraduate women to whom the complete MMPI had been administered, two groups of 20 *Ss* each were selected on the basis of extreme scores on Taylor's Manifest Anxiety Scale. Tachistoscopic recognition thresholds for two neutral practice words, four taboo words, and four neutral control words were determined. No significant relation between anxiety level and recognition threshold was found. Both groups showed a significant decline in thresholds for taboo words as a function of order of presentation, but the difference between taboo and neutral words was significant even in the final position. The magnitude of this residual difference was positively correlated with *Pd* score, but unrelated to *K*, *Hy*, *Sc*, or anxiety level. These results suggest that differences in threshold between taboo and neutral words (beyond those attributable to frequency of experience) can be understood in terms of differential readiness to report rather than perceptual distortion.

REFERENCES

1. FESTINGER, L. The significance of difference between means without reference to the frequency distribution function. *Psychometrika*, 1946, 11, 97–105.
2. HATHAWAY, S. R., & MCKINLEY, J. C. *The Minnesota Multiphasic Personality Inventory*. (Rev. Ed.) New York: Psychological Corporation, 1943.
3. HILGARD, E. R., JONES, L. V., & KAPLAN, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, 42, 94–99.
4. HOLTZMAN, W. H., CALVIN, A. D., & BITTERMAN, M. E. New evidence for the validity of Taylor's Manifest Anxiety Scale. *J. abnorm. soc. Psychol.*, 1952, 47, 853–854.
5. HOWES, D. H., & SOLOMON, R. L. A note on McGinnies' "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, 57, 229–234.
6. LUCAS, J. D. The interactive effects of anxiety, failure, and intra-serial duplication. *Amer. J. Psychol.*, 1952, 65, 59–66.
7. MEEHL, P. E., & HATHAWAY, S. R. The *K* factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *J. appl. Psychol.*, 1946, 30, 517–524.
8. MCGINNIES, E. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, 56, 244–251.
9. MCGINNIES, E. Discussion of Howes and Solomon's note on "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, 57, 235–240.
10. MCGINNIES, E., & SHERMAN, H. Generalization of perceptual defense. *J. abnorm. soc. Psychol.*, 1952, 47, 81–85.
11. ROSENBAUM, G. Stimulus generalization as a function of clinical anxiety. *Amer. Psychologist*, 1950, 5, 327. (Abstract)
12. SPENCE, K. W., & TAYLOR, JANET. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, 42, 183–188.
13. TAYLOR, JANET. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, 41, 81–92.
14. TAYLOR, JANET. A personality scale of manifest anxiety. *J. abnorm. soc. Psychol.*, 1953, 48, 285–290.
15. WILCOXON, F. *Some rapid approximate statistical procedures*. Stamford, Conn.: American Cyanamid Co., 1949.
16. WINNE, J. F. A scale of neuroticism: an adaptation of the Minnesota Multiphasic Personality Inventory. *J. clin. Psychol.*, 1951, 7, 117–122.

Received May 31, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

NEUROSIS AND TENSION: AN EXPLORATORY STUDY OF THE RELATIONSHIP OF PHYSIOLOGICAL AND RORSCHACH MEASURES[1]

JULIUS WISHNER
University of Pennsylvania[2]

THE major purpose of the present experiment is to explore some relationships between physiological and psychological measures of personality in such a way as to make possible the formulation of one or more general hypotheses concerning the psychophysiological functioning of the individual.

If personality is conceived as a unique pattern of more or less enduring behavioral tendencies, then it seems reasonable to assume that these should be measurable both psychologically and physiologically. The problem is what aspects of the pattern shall be measured. This depends, of course, upon what is perceived as basic in the understanding of the individual personality. While no precise predictions of the results were made in this study, nevertheless a particular kind of relationship was being sought: namely, the relation of the location and intensity of physiological activity to the direction[3] and intensity of psychological activity.

This seems to be a fundamental problem in the understanding of psychophysiological organization, for if the nature of this relationship were known, it would be a long step toward a fuller comprehension of an individual's mode of adjustment to his environment. Indeed, Duffy (6, 8) has proposed that intensity and direction of behavior are the only variables necessary for a description of personality. Even if one were to take issue with Duffy's omission of the functional significance of any particular behavior as an item to be considered in the understanding of an individual, the importance of the variables proposed would still be clear.

Two areas of study provide supporting background for the attempt to understand personality in terms of the intensity and direction of behavior. In the clinical field, the formulation that overt behavior may constitute a defensive system against unacceptable impulses is an almost universal one. Thus, for example, one may interpret overt passivity as an adjustive technique which prevents aggressive impulses from appearing at the surface.

This functional type of interpretation has been applied widely to the Rorschach test. Color (*C*) responses, for example, are presumed to indicate a tendency toward overt emotional reactivity, while a high percentage of purely form-determined (*F*%) responses is taken to indicate constriction. Other scoring categories can be interpreted along the same lines. Because this type of interpretation fits with the kinds of relationships sought in this study, and since various of the scoring categories have received validational support from experimental (15, 19) and statistical (20, 21) studies, it was decided to use the Rorschach test as a measure of the direction of behavior at the psychological level.

In the more experimentally oriented area, there is also evidence to suggest a relationship between overt activity and physiological measures. Many of these data have been reviewed by Duffy (8). Closely allied to the information sought in this experiment are the findings of Jones (9), Luria (14), Duffy (7), and Arnold (1). All of these studies indicate

[1] An abbreviation of this paper was presented at the American Psychological Association meetings in Chicago in 1951.

[2] This is a modified version of a dissertation submitted to the Graduate School of Northwestern University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer is indebted to Dr. Donald B. Lindsley for his patient criticism, encouragement, and supervision. Grateful acknowledgment is made to the Institute of Juvenile Research in Chicago and Dr. C. W. Darrow who made available the unique equipment of his laboratory, the Illinois Neuropsychiatric Institute which granted funds to pay some of the subjects, and Doctors W. Earley, S. Gamm, W. Obrist, J. Pathman, A. Rosenwald, and D. Shakow, who participated in various aspects of the research and made possible its completion.

[3] Direction of activity is here conceived of as a continuum extending from extreme inhibition of relevant motor activity, internalization of feelings, to extreme "acting out," impulsivity. This definition seems closely analogous to that offered by Duffy (6).

that it is the most overtly active individuals who show the least "involuntary" activity, and vice versa. Thus, the evidence points to the definite possibility of finding stable psychophysiological correlates of the direction of behavior. In order to explore such a possibility fully, it was decided to record simultaneously a number of physiological measures. In addition, two groups of subjects, normals (NR) and psychoneurotics (PN), were tested, since any useful psychophysiological index should be able to differentiate between these groups. If present thinking about psychopathology is at all close to the truth, NR and PN should differ as to behavioral direction and/or intensity.

Within such a context, then, the purpose of this study may be summarized as an attempt to find answers, however tentative, to the following questions:

1. Is there any correlation between Rorschach estimates of personality functioning and measures of physiological activity?

2. Will the direction of behavior or adjustment be similarly reflected in these measures?

3. Can a first approximation of an index of direction be constructed from these measures?

4. Will such an index differentiate neurotics from normals?

METHOD

Subjects

Eleven white male patients diagnosed by a psychiatrist as psychoneurotic, with anxiety as a major component, were utilized. These subjects (*Ss*) were obtained from two public clinics. A comparable group of 10 normals was examined. The criteria for "normality" were: (*a*) no history of referral to a psychiatrist, nor consideration given to self-referral; (*b*) no history of persistent illnesses of any kind; (*c*) no awareness of any pressing problems of adjustment, fairly well satisfied with life; (*d*) no history of any traumatic injuries, particularly to the skull. Some *Ss* in each group were paid for their time.

The mean age for the psychoneurotic (PN) group was 29.72 years, standard deviation 8.64, range 19–42. The mean age for the normal (NR) group was 25.10, standard deviation 6.77, range 18–39. The difference between the means is not significant. The mean age for the entire group of 21 *Ss* was 27.62, standard deviation 8.14, range 18–42. Group NR was probably slightly superior socioeconomically.

Procedure

Each *S* was given the Rorschach test which was administered and scored essentially as recommended by Beck (2). The following categories were calculated for each Rorschach protocol: *R*, *W*%, *D*%, *A*%, *P*, *T*/*R*, *F*+%, *F*%, *M*, *Sum C*, *Sum Y* (computed in the same way as *Sum C* for responses in which shading is a determinant), *Sum V* (computed in the same way as *Sum C* for responses in which perspective or distance is a determinant), and *S*.

Physiological measures were taken before and after the Rorschach, respectively, for about half of each group. Strict control of the timing of the two procedures was impossible owing to clinical exigencies. The following measures were recorded with *S* in a sitting position:

1. Muscle Action Potentials (MAP). (*a*) Frontal muscle above the eyebrows; bipolar, one electrode over each eyebrow (MAP–F). (*b*) Masseter muscle on right side of jaw; bipolar, electrodes about an inch apart (MAP–M).

2. EEG from the left side of the skull; only the left occipital monopolar lead was used in the final analysis.

3. Eye movement, from electrodes placed between the eyebrows and to the left of the left eye, respectively.

4. EKG, from electrodes attached to each upper arm.

5. Respiration, from pneumograph attachments over the chest and stomach.

6. GSR, from electrodes on the left palm and wrist.

7. Relative blood pressure, from a cuff above the left ankle, inflated to a pressure between systolic and diastolic, as indicated by a manometer.

The procedure was as follows:

1. A recording during the normal resting state was made for approximately two minutes.

2. The following questions were asked of *S*, at approximately 15-second intervals after *S*'s reply: (*a*) How much is two times two? (*b*) How much is three times three? (*c*) How much is four times four? This was sometimes varied to: How much is five times four? (*d*) How much is 13 times 64? A recording of about two minutes' duration was made after the last question. If no answer was given, *S* was told to forget about the problem, and that no further questions would be asked.

3. The cold pressor (CP) test was then administered as follows: The experimenter (*E*) placed *S*'s right hand in a pitcher of iced water. *S* kept his hand in the water for 30 to 60 seconds. (It was necessary to vary the timing of this procedure because several *Ss* complained that it was too painful.) *E* then removed *S*'s hand from the water and returned it to its previous position, with instructions to let it dry by itself. Recording was continued throughout.

4. When physiological recovery from the CP test seemed apparent, *S* was instructed to hyperventilate. He was told to stop after 40 inhale-exhale cycles.

5. Following recovery, a loud blast on a horn was blown.

These procedures were intended to sample several forms of reactions. In the case of the questions in item 2, a particular mental set was broken with the intention of producing emotional arousal; items 3

and 4 provided different examples of physiological stress, while the startle response was sampled in item 5.

Apparatus

The *S* was seated in a darkened and electrically shielded room. He was grounded by a clip electrode attached to one ear. All recordings were made in an adjoining room, from which *E* could watch *S* through an opening.

MAP and EEG were recorded on an 11-channel electroencephalograph, composed of two Offner units: Type A (6 channels) and Type B (4 channels), plus one additional channel. The pen-writing unit consisted of Rochelle salt crystals operating on 240 volts.

GSR, respiration, and relative blood pressure were recorded on a Darrow Behavior Research Photopolygraph previously described in the literature (4).

Eye movement was recorded on the EEG paper by an electromagnetically driven pen, operated by a separate set of amplifiers. Respiration was recorded on the EEG paper by a pneumatically driven pen, as well as on the photopolygraph.

Small solder discs, 3 mm. in diameter, constituted the electrodes for EEG, MAP, and eye movement. The "Darrow electrodes" (5) were used for GSR. EKG was recorded from electrodes consisting of gauze-covered wire mesh soaked in a sodium chloride solution.

Analysis of the Records

The following measures were calculated before and after each stimulus condition: 1. EEG: The record was analyzed for frequency, amplitude, and per cent time alpha essentially as described by Lindsley and Cutts (13). 2. MAP: Kennedy and Travis (10, 11) have found that amplitude of MAP–F yields reliable individual differences. To simplify the task of measurement, the trough to peak height of single representative spikes was measured at previously determined time intervals within each portion of the record to be analyzed. A minimum of 3 values and a maximum of 10 were determined for each portion of the record. Thus, a measurement was taken at intervals of from every 5 to every 10 sec. All values for each portion of the record were averaged, and this mean was used as the representative value. Measurements were made with a transparent plastic ruler to the nearest .5 mm.; for most records, this meant to the nearest 3 μv. Unlike the frontal muscles, the masseter muscle had no consistent and stable amplitude. Therefore, for MAP–M, it was decided to count the number of "contractions" of the muscle for each portion of the record analyzed. The criteria for a contraction were the presence of five or more spikes of at least 30 μv., and a total duration of the activity for at least .5 sec. In the case of prolonged activity, a value of 1 was assigned to each 2-sec. duration. 3. GSR: Lacey and Siegel (12) have shown that change in conductance is an adequate measure for statistical computation. This value expressed in micromhos was used in the present study. 4. Respiration rate (Resp.): The number of inspiration-expiration cycles / minute was calculated for each relevant portion of the record. 5. EKG: Heart rate / 10 sec. (HR) was counted at the appropriate points. When more than one value was obtained (if the portion of the record analyzed was greater than 10 sec.), the mean was taken. 6. Eye movement and relative blood pressure were utilized as qualitative controls on the EEG, and were not themselves quantified.

RESULTS

The Rorschach Test

None of the differences between the means of the Rorschach scoring categories is statistically significant. The PN group tends to be higher in T/R and $F\%$, indicating that this group tends to be slower in its responses and is less sensitive to the various nuances in the test, such as color and shading.[4] These are the only differences for which t is greater than 1.

The Physiological Measures

MAP–F differentiates significantly between the groups in 7 of the 14 portions of the record analyzed; the greatest absolute difference occurred during CP. At the end of CP, the mean microvoltage for the PN group was 16.53, for the NR group, 26.69. The PN group is significantly higher in HR in 10 portions of the record, with the greatest differences again occurring during and after the CP test. This is strikingly similar to the results of White and Gildea (18). No other significant differences between the groups were found.[5]

Several of the other measures show consistent differences in favor of one group or the other. Although these differences are not statistically significant, they merit consideration because they all approach the .05 level. The PN group is consistently higher in Resp. and % time alpha, and lower in GSR. The difference between the means of the two groups during initial resting respiration rate falls just short of the .05 level of confidence.

The groups tend to maintain the same relationship to each other for all portions of the

[4] For tables supplementary to this article order Document 3724 from ADI Auxiliary Publications, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting $1.00 for microfilm (images 1 in. high on standard 35-mm. motion picture film) or $1.00 for photostats readable without optical aid.

[5] See footnote 4.

records analyzed. The correlations between the resting level and the mean of all conditions for MAP–F, HR, and Resp. were .69, .89, and .86 respectively. It seemed justifiable, therefore, to compute a mean for each measure, combining all conditions under which it was used. These means of the two groups are significantly different only in MAP–F and HR.[6]

Relations between Rorschach and Physiological Measures

With both groups treated as a whole, each physiological measure was correlated with each Rorschach category. These coefficients are given in Table 1.

If we extract all of the significant coefficients, it is seen that there are significant positive correlations among the following: MAP–F and *Sum C*, HR and *F*+%, GSR and *Sum Y*, GSR and *Sum C*, Resp. and *A*%, *Resp.* and *T/R*, Resp. and *F*%; the correlation of MAP–M and *F*% is almost significant at the .05 level. Significant negative correlations are: MAP–M and *Sum C*, GSR and *A*%; again, the negative correlation of MAP–M and *M* is just below that required at the .05 level. All of these intercorrelations were computed separately for each group in order to determine whether there might be differences between the groups. Only HR and *F*+% showed a large difference between the groups, and, as a result, both of these scores were excluded from further consideration. All of the other significant correlations for both groups combined showed very small, unreliable differences between the groups.

Close scrutiny of these significant coefficients indicates that some pattern does exist. Thus, MAP–F and GSR have in common significant positive relationships with *Sum C*. MAP–M and Resp. have in common positive relationships with *F*%. Moreover, MAP–M and Resp. tend to have the same signs for their coefficients with the Rorschach scores; similarly MAP–F and GSR. It is particularly interesting that MAP–M and Resp. have negative correlations with *Sum Y* (which has a significant positive correlation with GSR), *Sum C* (whose significant positive relationship with both GSR and MAP–F has been noted), and *M*; on the other hand, MAP–F and GSR have negative relationships with *A*% (which is positively related to MAP–M and Resp.), *T/R* (significantly related to Resp.), and *F*% (again positively related to MAP–M and Resp.).

TABLE 1

CORRELATIONS OF PHYSIOLOGICAL AND RORSCHACH MEASURES

(N=21)

RORSCHACH	PHYSIOLOGICAL MEASURES							
	MAP–F	MAP–M	HR	GSR	RESP.‡	EEG FREQ.	EEG AMP.	EEG % TIME
R	−.05	−.37	.04	.37	−.31	.24	.16	−.02
W%	.01	.33	.07	−.32	.17	−.24	−.11	.08
D%	.02	−.15	−.17	.29	−.29	.20	.09	−.15
A%	−.01	.29	.09	−.45*	.52*	−.28	.04	.15
P	−.04	−.14	.05	.20	−.25	−.11	−.11	−.06
T/R	−.32	.02	.12	−.35	.54**	−.35	.20	.32
F+%	−.10	.32	.46*	−.03	−.10	.13	.17	.39
F%	−.03	.41	.11	−.42	.50*	.16	−.10	−.04
M	−.02	−.41	.02	.29	−.16	.14	−.17	−.04
Sum Y	.00	−.25	−.17	.59**	−.34	−.02	.28	.18
Sum V	.03	−.16	−.08	.26	−.33	−.24	.11	−.02
S	−.04	−.18	−.01	.25	−.04	−.24	−.05	−.06
Sum C	.45*	−.48*	−.08	.56**	−.03	.01	.18	.06

‡ N=20.
* Significant at .05 level of confidence.
** Significant at .01 level of confidence.

[6] See footnote 4.

It seems likely, therefore, that two *groups* of scores can be extracted from both the physiological and Rorschach measures. MAP–M and Resp. appear to go together with *A*%, *T/R* and *F*%; MAP–F and GSR go together with *M*, *Sum Y*, and *Sum C*. Moreover, in view of the general tendency toward negative relationships between these two groups, it seems probable that they represent opposing behavior tendencies.

In order to determine whether a single index of the direction of the behavior could be constructed, the following procedure was used: each of the scores mentioned in the paragraph above was converted into a standard (z) score. They were then added together within the groups outlined, with the Rorschach and physiological data treated separately. Thus, Group I Physiological contained the sum of the z scores of MAP–M and Resp.; Group I Rorschach was the sum of the z scores of *A*%, *T/R* and *F*%; Group II Physiological was the sum of the z scores of MAP–F and GSR; Group II Rorschach was the sum of the z scores of *M*, *Sum Y*, and *Sum C*. In addition, for the Rorschach and physiological scores separately, the sum of the z scores of Group II was subtracted from the sum of the z scores of Group I, so that a single index for each *S* in the experiment was constructed. The intercorrelations of these groupings are shown in Table 2. It will be seen that the Rorschach and physiological scores resulting from the subtraction of Group II from Group I have a correlation coefficient of .66. If the procedure outlined is valid, it is apparent that the direction of behavior can be similarly described at the levels of both the physiological and Rorschach measures. It seems probable, however, that most of this correlation is contributed by the relatively high correlation existing between the Group I scores of the Rorschach and physiological data.

The grouped scores, i.e., the sum of the z scores as indicated above, were computed for each *S* and for each diagnostic group.[7] The physiological measures show consistent differences in the direction of the two groups. Group NR tends to have higher GSR+MAP–F, while Group PN shows higher Resp.+MAP–M. The Rorschach measures indicate consistently higher scores (with signs disregarded) for Group PN. Thus, the neurotics' Rorschachs indicate either excessive overcontrol, i.e., high (*A*%+*T/R*+*F*%), or excessive lack of control, ie., records dominated by (*M*+*C*+*Y*). The normals, on the other hand, show, in general, protocols balanced between the two groups of scoring categories. These differences become more impressive when it is noted that Group NR contains at least one schizoid and one severe neurotic (as indicated by clinical impression and independent ratings of the Rorschach protocols) who, nevertheless, met the criteria of normality established for this study.

TABLE 2

INTERCORRELATIONS OF SELECTED STANDARD SCORES

(N=20)

RORSCHACH	PHYSIOLOGICAL I (MAP–M+RESP.)	PHYSIOLOGICAL II (MAP–F+GSR)	I–II
I (*A*%+*T/R*+*F*%)	.71	−.43	.72
II (*M*+*Sum Y*+*Sum C*)	−.60	.31	−.57
I–II	.68	−.38	.66

DISCUSSION

In terms of personality characteristics, what could be the source of the moderately high

[7] See footnote 4.

correlations between ($A\%+T/R+F\%$) and (MAP–M+Resp.), and between [($A\%+T/R+F\%$)−($M+Sum\ Y+Sum\ C$)] and [(MAP–M+Resp.)−(MAP–F+GSR)]? Clearly, no differentiation between these groups of scores on the physiological side can be made on the basis of heightened or lowered autonomic activity. The fact is that GSR and Resp. are both, at least in part, measures of autonomic activity, and, as can be seen in Table 1, they tend to be correlated with the Rorschach measures in precisely opposite ways, i.e., the signs of their respective correlations with any particular Rorschach score are reversed. As a matter of fact, GSR and Resp. tend to be inversely correlated, although not significantly so in this study.

Consideration of the Rorschach scores involved may yield a better understanding of the meanings of the correlations as well as the source of the differences between groups PN and NR. We may note that $A\%$, T/R, and $F\%$ all tend to indicate low sensitivity and/or reactivity to the more subtle nuances of the Rorschach test. On the other hand, *M*, *Sum Y*, and *Sum C* show sensitivity to the color, shading, and movement aspects of the test. These latter scores tend to be negatively correlated with MAP–M and Resp., and positively correlated with MAP–F and GSR. Darrow (3) has reviewed the literature on the GSR and concludes that it is functionally related to anticipatory behavior. Travis and Kennedy (16, 17) have shown that MAP–F is positively correlated with alertness in tracking behavior. Thus, it appears that the *Ss* who remained alert, waiting for the next stimulus during the physiological tests, were also the ones who, in general, were the most sensitive individuals on the Rorschach test.

If we examine the nature of the measurement of MAP–M in this study, it will be clear that it probably reflects a diffuse discharge of tension. It will be remembered that MAP–M was measured in terms of the number of "contractions," which would result only from restless or "involuntary" tensing of the jaw muscles or from swallowing. Such behavior is a common characteristic of "nervous" or anxious individuals. Similarly, the tendency to breathe quickly and shallowly which results in an increased respiration rate is characteristic of many psychologically disturbed individuals. That MAP–M+Resp. is positively correlated with $A\%+T/R+F\%$ on the Rorschach test is therefore not surprising. In the discussion to follow, the restless activity indicated by MAP–M and Resp. will be labelled "diffuse" (D) activity; the anticipatory reactions indicated by MAP–F and GSR will be called "focused" (F) activity.

In terms of personality characteristics, then the analysis above may be summarized as follows: the lower the sensitivity of an individual on the Rorschach test, the greater the D activity. Moreover, group PN tends to be greater in D activity than group NR. In terms of a single test, it may be said that the greater the preponderance of D over F activity, the more likely is the individual to be psychologically disturbed, as shown by the Rorschach test.

From this, we may construct a general hypothesis: Psychopathology is characterized by the preponderance of diffuse activity when compared to the amount of focused activity. The leap from psychoneurotic patients with anxiety as a major symptom to psychopathology in general is justified only by the attempt to generalize the hypothesis, and by the commonly accepted postulate that anxiety is the central core of psychopathology. In the present study, a single number was arrived at by the formula D—F. This yields, in the sign of the result, the dominant direction of an individual's activity.

It is obviously not necessary to test this hypothesis only through physiological measurements. It is stated broadly enough so that it can be tested in relation to any task which can be analyzed in terms of the activity appropriate to it. Even if the hypothesis is tested by physiological measurements, it should be noted that no single physiological test of D or F activity is here proposed. For example, if MAP–F were measured in the same units as MAP–M, it is likely that it would then become an indicator of D activity. Similarly, if GSR's occur in the absence of objective stimuli, then these too could be considered D activity. Thus, it is not the particular physiological activity measured which is important. Rather, the measure must be analyzed in each case in relation to the conditions under which it is taken. Under

onditions where *S* can do nothing but wait nd remain alert, as in this study, and when easurements at the resting level are made, he measures utilized should be adequate, lthough they undoubtedly can be improved.

It seems apparent, therefore, that all of the uestions presented in the introduction may e answered affirmatively. However, it may e well to emphasize that the explanations nd hypothesis offered are of the *post hoc* ariety, and that no stable conclusions can be rawn without repetition and elaboration of his study. At the same time, it is of interest hat the work of Luria (14), Duffy (7), and Arnold (1), previously cited, may all be interpreted in terms of the hypothesis proposed. Whether it can lead to a more clearly defined dimension of psychological health or psychopathology, and therefore to more fruitful research in this field, only further investigation can show.

Finally, some lessons may be gleaned from the results of the various experimental stimuli utilized. The high intercorrelations between the responses to the stimuli indicate that it is unnecessary to present complex stimuli in order to elicit reactions in which constant individual differences are maintained. It is clear that the groups maintained the same relative position on almost all of the physiological measures regardless of the stimulus presented. Of course, in order to elicit the GSR some stimulus is indispensable. On the basis of the results gathered in this study, it appears that the cold pressor test and hyperventilation tend to maximize the differences between the groups, as well as the reactions of any given individual. Thus, further research of this kind could be made much more economical by utilizing only one or two stimulus conditions.

SUMMARY

The purpose of this study was to explore the possible relationships between physiological and Rorschach measures of personality characteristics in order to construct one or more general, testable hypotheses. The particular kinds of data sought were in the area of the relationship between overt and physiological activity. It was felt that if stable psychophysiological correlates of the direction of an individual's behavior could be found, it would increase the possibility of fruitful research in this area. The specific questions posed at the beginning of the study were: 1. Is there any correlation between Rorschach estimates of personality functioning and measures of physiological activity? 2. Will the direction of behavior or adjustment be similarly reflected in these measures? 3. Can a first approximation of an index of direction be constructed from these measures? 4. Will such an index differentiate between neurotics and normals?

Eleven white, male psychoneurotic patients, with anxiety as a major component, and ten normals were tested with the Rorschach test, and on the same day, measurements of their physiological activity were made under various stimulus conditions. These measures included electromyograms from the frontal and masseter muscles, electroencephalograms, heart rate, respiration rate, and the galvanic skin response.

None of the Rorschach scoring categories, considered singly, differentiated between the two groups. Heart rate was significantly higher for the neurotic group, and respiration rate also tended in this direction. The normals had higher galvanic skin responses and muscle action potentials from the frontal muscles.

Analysis of the correlations between the Rorschach and physiological variables indicated that two *groups* of scores could be extracted. The sum of the *z* scores of *A%*, *T/R* and *F%* correlated with the sum of the *z* scores of respiration rate and the number of "contractions" of the masseter muscle. *M*, *Sum Y*, and *Sum C*, added together, correlated positively with the sum of the *z* scores of the muscle action potentials from the frontal muscle and the galvanic skin response.

It was hypothesized that the first of these groups of scores represented diffuse (D) discharges of tension; on the other hand, the latter group was thought to represent focused (F), anticipatory behavior. The formula D—F, therefore, should yield a single index of the direction of behavior, in terms of the preponderance, or absence, of D activity for any given individual. D—F computed separately for the Rorschach and physiological measures yielded a correlation coefficient of .66.

It is felt that these results necessitate affirmative answers to the questions posed. The general hypothesis that psychopathology is characterized by a preponderance of diffuse, unfocused activity is proposed for further testing.

REFERENCES

1. ARNOLD, M. B. A study of tension in relation to breakdown. *J. gen. Psychol.*, 1942, **26**, 316–346.
2. BECK, S. J. *Rorschach's test.* Vol. I. *Basic processes.* (2nd Ed.) New York: Grune and Stratton, 1949.
3. DARROW, C. W. Neural mechanisms controlling the palmar galvanic skin reflex and palmar sweating. *Arch. Neural. Psychiat.*, 1937, **37**, 641–663.
4. DARROW, C. W. The behavior research photopolygraph. *J. gen. Psychol.*, 1932, **7**, 215–219.
5. DARROW, C. W. Uniform current for continuous standard unit resistance records. *J. gen. Psychol.*, 1932, **6**, 471–474.
6. DUFFY, E. A systematic framework for the description of personality. *J. abnorm. soc. Psychol.*, 1949, **44**, 175–190.
7. DUFFY, E. Tensions and emotional factors in reaction. *Genet. Psychol. Monogr.*, 1930, **7**, 1–79.
8. DUFFY, E. The concept of energy mobilization. *Psychol. Rev.*, 1951, **58**, 30–40.
9. JONES, H. E. The study of patterns of emotional expression. In M. Reymert (Ed.), *Feelings and emotions.* New York: McGraw-Hill, 1950. Pp. 161–168.
10. KENNEDY, J. L., & TRAVIS, R. C. Prediction and automatic control of alertness. II. Continuous tracking. *J. comp. physiol. Psychol.*, 1948, **41**, 203–210.
11. KENNEDY, J. L., & TRAVIS, R. C. Prediction of speed of performance by muscle action potentials. *Science*, 1947, **105**, 410–411.
12. LACEY, O. L., & SIEGEL, P. S. An analysis of the unit of measurement of the galvanic skin response. *J. exp. Psychol.*, 1949, **39**, 122–127.
13. LINDSLEY, D. B., & CUTTS, K. K. Electroencephalograms of "constitutionally inferior" and behavior problem children. *Arch. Neurol. Psychiat.*, 1940, **44**, 1199–1212.
14. LURIA, A. R. *The nature of human conflicts.* New York: Liveright, 1932.
15. SIIPOLA, E. M. The influence of color on reactions to ink blots. *J. Pers.*, 1950, **18**, 358–382.
16. TRAVIS, R. C., & KENNEDY, J. L. Prediction and automatic control of alertness. I. Control of lookout alertness. *J. comp. physiol. Psychol.*, 1947, **40**, 457–461.
17. TRAVIS, R. C., & KENNEDY, J. L. Prediction and control of alertness. III. Calibration of the alertness indicator and further results. *J. comp. physiol. Psychol.*, 1949, **42**, 45–57.
18. WHITE, B. V., & GILDEA, R. F. "Cold pressor test" in tension and anxiety: a cardiochronographic study. *Arch. Neurol. Psychiat.*, 1937, **38**, 954–984.
19. WILLIAMS, M. An experimental study of intellectual control under stress and associated Rorschach factors. *J. consult. Psychol.*, 1947, **11**, 21–29.
20. WITTENBORN, J. R. A factor analysis of Rorschach scoring categories. *J. consult. Psychol.*, 1950, **14**, 261–267.
21. WITTENBORN, J. R. Statistical tests of certain Rorschach assumptions: the internal consistency of certain scoring categories. *J. consult. Psychol.*, 1950, **14**, 1–19.

Received May 26, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

ON THE THEORY OF PSYCHOANALYTIC DISPLACEMENT[1]

ROBERT R. BUSH AND JOHN W. M. WHITING
Harvard University

DURING the past 10 years, there have been a number of studies, theoretical and experimental, which have related psychoanalytic concepts to stimulus-response behavior theories and related experiments. In 1943, Sears (14) reviewed such studies, but since then much important work has been done. Perhaps the most extensive contribution is that of John Dollard and Neal E. Miller (5). Among the Freudian mechanisms which have been studied in some detail is the phenomenon of displacement. In an important paper, Miller (12) presented a theory to explain displacement in terms of stimulus generalization. As a direct result of this paper, at least three sets of experiments have been completed (1, 13). In addition, Miller's theory has had considerable influence on the interpretation of doll-play data in the Laboratory of Human Development at Harvard and of cross-cultural data by Whiting and Child (16). That theory, provocative as it has been, is somewhat unsatisfactory and incomplete. This paper is an attempt to remedy these shortcomings.

We do not remedy all of the shortcomings of Miller's model, however. Any formal quantitative model of a complex psychological process necessarily involves numerous abstractions and simplifying assumptions and as a result is inadequate for describing many, indeed most, instances of that process. The price of precision is the loss of generality and many psychologists may feel that the price is often too great. The argument hinges upon, it seems to us, the strategy one uses in his research, i.e., whether precision or generality has priority. There is always hope of making a general theory more precise and of making precise theory more general, but which is easier is a matter of faith not fact. Our faith, and apparently Miller's also, is in the procedure of slowly making more general a model with built-in precision. When simplifying assumptions are made, attention is focused on the factors which make the phenomena complex. Generality is increased by examining these assumptions, one by one, and finding suitable substitutes. Such a mechanism for improvement is hard to find for increasing the precision of a general theory.

The original concept of displacement was suggested by Freud. In his writings, the displacement mechanism appears mainly in relation to the analysis of dreams and wit, but at one point at least (9), Freud says that as a result of displacement ". . . the lonely spinster transfers her affection to animals, the bachelor becomes a passionate collector, the soldier defends a scrap of colored cloth—his flag—with his life-blood, in a love affair a clasp of the hands a moment longer than usual evokes a sensation of bliss, or in *Othello* a lost handkerchief causes an outburst of rage . . ." (9, p. 246). In *Frustration and Aggression* (4), the displacement of aggression is discussed in some detail. The basic idea of the displacement concept is that when a particular kind of behavior such as aggression is inhibited by punishment or fear, it will appear in different forms or toward objects other than the ones originally involved.

Miller's Theory

In his theoretical paper, Miller related the psychoanalytic concept of displacement to the stimulus-response process of generalization (12). He derived a displacement effect from a set of assumptions about how approach and avoidance responses generalize from one stimulus situation to other similar situations, when the approach or positive response was inhibited by a conflicting response. He lists five basic assumptions as follows:

> . . . (1) that the direct response to the original stimulus generalizes to other similar stimuli, with the amount of generalization becoming smaller the less similar the stimuli; (2) that the response which conflicts with the occurrence of the direct response

[1] This research was aided by the Laboratory of Human Development and the Laboratory of Social Relations of Harvard University. During the past two years the authors had numerous helpful discussions with F. R. Brush, I. L. Child, H. Faigin, E. Hollenberg, W. O. Jenkins, W. F. John, N. E. Miller, F. Mosteller, P. S. Sears, and R. R. Sears. Their contribution to our thinking is gratefully acknowledged.

to the original stimulus also generalizes to other similar stimuli, becoming weaker the less similar the stimuli; (3) that the gradient of generalization of the conflicting response falls off more steeply with dissimilarity than does that of the original response which it inhibits; (4) that when two or more incompatible responses are simultaneously excited, the one with the greatest net strength will be the one which will occur; and (5) that the net strength of a response will be its strength minus that of any response incompatible with it which is excited at the same time (12, pp. 167–168).

From these assumptions, Miller makes a number of interesting deductions, some of which we will discuss later, but at this point we will examine his assumptions and point out some difficulties which they generate.

The chief problem raised by Miller's assumptions is embodied in his fifth assumption in which he proposes that the net strength of one of two conflicting habits can be obtained by subtracting the strength of the other habit from it. Although we agree with the underlying principle that incompatible habits in conflict tend to interfere in some manner with one another, we do not feel that algebraic summation is the most appropriate hypothesis for calculating this interference. What sort of mathematical assumption one makes, however, depends in large part on the operational measure of response strength which is chosen. If that measure is amplitude (e.g., strength of pull in grams), as was the case in Miller's original work on conflict (11), algebraic summation seems clearly appropriate. However, our model is chiefly concerned with probability measures of habit strength where the laws of combination of probabilities must be examined in detail. The experiments on displacement, in contrast to those on conflict, deal with probabilities and rates rather than amplitudes.

However, Miller is not clear that he is indeed using amplitude as his basic measure. In fact, on the one hand, his fourth assumption—that the response with the greatest net strength will occur—suggests that he is using prepotency as an index of habit strength, and, on the other hand, in his experiments on displacement he has used not only amplitude (volume of water drunk), but also rate (speed of running) and probability (number of errors, proportion of animals). We feel that these ambiguities must be eliminated in order to arrive at a satisfactory model for displacement.

The third assumption—that the conflicting responses decrease more rapidly in strength than the direct response as similarity is decreased—seems to us to be unnecessarily restrictive. In some cases this assumption may make sense, but it is easy to think of counterexamples. As Miller himself suggested to us in conversation, it is possible to alter the cues associated primarily with the direct response without appreciably changing the cues connected with the conflicting responses; if this were done we might reasonably expect the opposite of Miller's third assumption to be the case. We shall have more to say about this point later.

We do not wish to take the position that precise, quantitative, operational definitions are always necessary. Nevertheless, when definitions can be sharpened they are of tremendous aid in building a quantitative theory. We hope that it is evident that it is Miller's assumptions and definitions, not his conclusions, with which we quarrel. Even then, we have often felt that our concern was too much for detail and not enough for broader issues and implications. In any case, this paper is concerned with detail and it is our hope that we will be one step closer to a precise theory of displacement phenomena.

The Set-Theoretic Model

The mathematical model for analyzing displacement phenomena which we will present is based upon a set-theoretic model for stimulus generalization presented by Bush and Mosteller (3). The basic features of this latter model will be reviewed here. A given stimulus situation is described by a set of stimulus elements denoted S. (Although we do not exclude response produced stimuli as stimulus elements in our model, we have not explicitly considered problems peculiar to response mediated generalization.) A second stimulus situation, similar to the first, is denoted by S'. The intersection of S and S', i.e., the subset of elements common to S and S', is labeled I. An index of similarity, η, of S' to S is defined by:

$$\eta = \frac{m(I)}{m(S)} \qquad [1]$$

where the notation m ( ) is used to denote the *measure* of the set or subset named in the parenthesis; this measure would correspond to number of elements if the elements were

of equal importance to the organism, but in general the measure of a set of elements indicates the relative importance of that set.

Sets of elements and their measures are not operationally defined and so the similarity index, η, is not operationally defined by equation [1]. This index is given operational meaning, however, by the following considerations. Let all the elements of S be "conditioned" to a particular response r; the probability p that r will occur in S is thus unity. (Fundamentally, our postulate is the one given by Bush and Mosteller (3): Stimulus elements exist in one of several "states" and each state is formally defined by the conditioned generalization. We cannot define amount of generalization in terms of the degree of similarity without being circular, but we will have no need for such a definition of the amount of generalization. In other words, the relationship between amount of generalization and degree of similarity is not a *law* but a *definition* of the degree of similarity. This is analogous to the modern view originally clarified by Ernst Mach, that Newton's second "law" is not a law at all but a definition of force as the product of mass and acceleration. Furthermore, definition [2] is not relative to the particular response r which was chosen since we have demanded

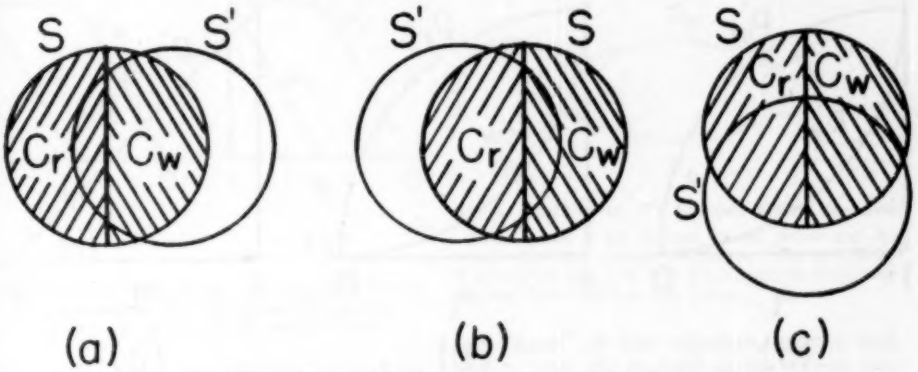

FIG. 1. THREE WAYS IN WHICH A SET S' CAN BE SIMILAR TO A SET S

The portion C_r of S has been conditioned to response r and the portion C_w of S has been conditioned to response w. In (a), S' contains more of C_w than of C_r; in (b), S' contains more of C_r than of C_w; and in (c) S' contains an equal amount of C_r and C_w.

responding class of responses. Thus we speak of an element as conditioned to one response class or another.) The similar set S' will now contain elements conditioned to r—those contained in I—as well as elements not conditioned to r—those in S' but not in I. We assume that no prior conditioning of r in S' has taken place so that no elements in S' and not in I are conditioned to r. The probability, p', that r will occur in S' is then, by definition, the measure of the elements in S' conditioned to r divided by the measure of all the elements in S'. Hence we have:

$$\eta = p' \text{ (when } p=1). \qquad [2]$$

This equation provides the necessary operational definition of η, since p' can be measured (in principle, at least) for any set S' after conditioning has been completed in S. Thus, we have operationally defined degree of similarity in terms of the amount of obtained generalization.

that S be completely conditioned to r and that there be no prior conditioning of r in S'. However, S and S' must be such that response r is possible. For example, if r is bar-pressing, the bar must be present in both S and S'. The set of alternative responses available to the organism must be identical in S and S'.

We now depart from the assumption of equal proportions made by Bush and Mosteller (3) and consider a set S in which some portion C has been conditioned to response r. A second set S' may be "similar" to S in various ways for a given index of similarity: (a) the intersection may contain mostly elements not conditioned to r, and so $p' < \eta p$, or (b) the intersection may consist mainly of elements conditioned to r, and so $p' > \eta p$, or (c) the intersection may be such that $p' = \eta p$. (See Fig. 1.)

The three cases of similarity just described correspond, we believe, to common sense notions about similarity. To illustrate at an intuitive level, let us consider a child who has learned to roll a large round ball, say a basketball (S in our example). Let us further assume that he has been more conditioned to its roundness than to its size. Then let us consider how he might react to three new objects (S' in our example), all of which are somewhat similar to the basketball (S). Case (a), which could involve changing its shape more than its size, might be represented by a large cubical block. Case (b), which would involve changing its size more than its shape might be represented by a tennis ball. Case (c) which would involve changing both its size and its shape about equally might be represented by a football. Common sense would say and our model would predict, first that he will have somewhat less tendency to roll any of the new objects than to roll the old familiar basketball, and second, that the order of probability of his rolling the new objects would be tennis ball most, football next, and large block least.

We are now in a position to describe what happens in a conflict situation. In a stimulus population S, an organism has been both rewarded and punished for making a response r. Hence, we assume that part of S, called C_r, is conditioned to r while another part of S called C_w, is conditioned to withdrawal from the situation. This view of the effects of punishment seems to be now generally accepted; punishment causes withdrawal responses to be reinforced since withdrawal causes a cessation of pain or anxiety. As an illustration, let half of S be conditioned to r, and half conditioned to withdrawal responses w, as shown in Fig. 1. For simplicity, we assume that all elements of S are conditioned to either r or w. Now a situation S', similar to S by an amount η may contain (a) less of C_r than C_w, may contain (b) more of C_r than of C_w, or may contain (c) an equal portion of C_r and C_w. Since we must have $p'_r+p'_w=\eta$, where p'_r and p'_w are the probabilities of responses r and w, respectively, these three cases lead to curves of generalization of the type indicated in Fig. 2. Miller (12) concludes that only case (b) occurs in empirical phenomena. Limiting cases of type (a) and (b) are diagramed in Fig. 3. In Fig. 3(a), p'_w remains constant until $p'_r=O$, while in Fig. 3(b) the reverse is true. We have $p'_r+p'_w=\eta$ for all values of η.

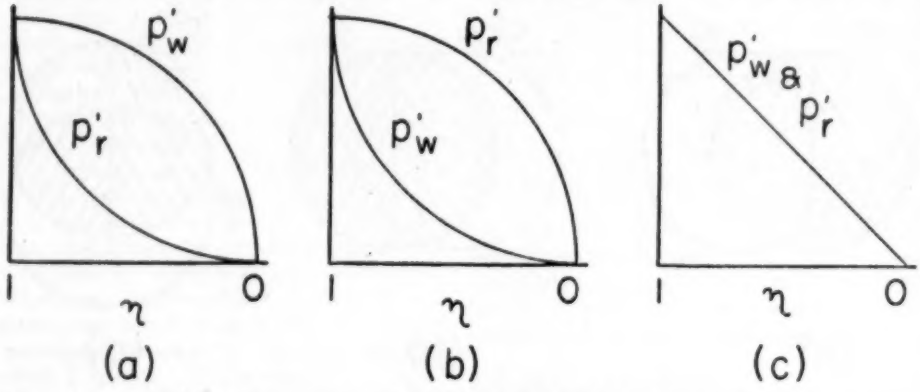

FIG. 2. ILLUSTRATION OF THE DEPENDENCE OF p_r' AND p_w', THE PROBABILITIES OF RESPONSES r AND w, RESPECTIVELY, IN S', ON THE SIMILARITY INDEX η. THE THREE CASES CORRESPOND TO THE CASES DEPICTED IN FIG. 1.

A major point which emerges from the model at this stage is that the probability p'_r of response r in any S' can never be greater than the probability p_r of response r in S. There will always be a generalization decrement. This deduction was made earlier by Bush and Mosteller (3). It might seem therefore that displacement—defined as an *increase* in response strength along a similarity dimension—cannot be predicted by the model. But this is not quite so, as we shall see shortly. The conclusion reached in this regard depends in a crucial way upon the measure of response strength employed.

Opportunities for Responding

The several probabilities discussed above were defined to be the probabilities of various responses in situation S or S', i.e., the probabilities of various responses *given an opportunity for making those responses.* We now inquire about the frequency of opportunities for making a choice among the alternative responses. What is peculiar about problems in conflict, or more generally about problems involving punishment, is that the withdrawal responses which get established *remove the organism from the situation.* As a result, the organism has fewer opportunities for making the response. Examples from human behavior are numerous: anxiety about sex may cause a person to avoid sexually stimulating situations, fear of people leads a man into solitude, punishment of aggression develops a distaste for situations which might lead to frustration, etc. Indeed, this phenomenon is related to a central problem in psychotherapy. To sharpen up our picture we will give an example of an experimental problem. In the paper by Brush, *et al.* (1), the following experiment is reported. A pigeon is allowed to peck an illuminated window in a Skinner box. Part of the time the pecking response is followed by food and other times by an electric shock through the bird's feet. The onset of shock led to various reflexes and "anxiety behavior," such as wing-flapping, thrashing about the cage, etc. According to modern theories of punishment, those responses which occurred when the shock stopped were reinforced. From our point of view, the crucial point is that all those responses strengthened by shock caused the bird to be in such a position in the cage that it could not peck the illuminated window immediately. With respect to the pecking response and the stimulus situation S in the immediate neighborhood of the window, the responses induced by shock amount to withdrawal responses. As a result, the bird spent a smaller fraction of the experimental time in the immediate neighborhood of the window where it could peck than that bird did previous to being punished. Thus, the bird had fewer "opportunities" to peck the window.

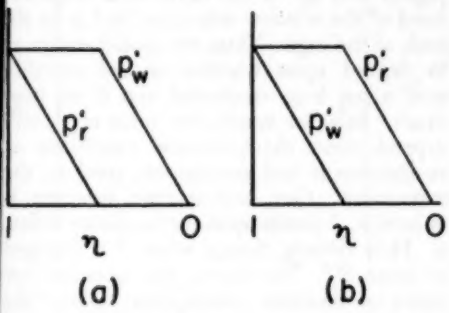

FIG. 3. LIMITING CASES OF FIG. 2(*a*) AND 2(*b*). IN (*a*), p_r' DECREASES TO ZERO BEFORE p_w' BEGINS DECREASING AND IN (*b*) THE REVERSE IS TRUE.

We shall formalize and abstract these arguments in the following way. Let there be two "states" of the organism, state 1 and state 2. For the example of the pigeon just mentioned, state 1 corresponds to being in situation S, the immediate neighborhood of the window, and state 2 corresponds to being elsewhere, in the back of the cage for example. For analytic purposes, time will be quantized into intervals of length h and each such interval will be called a "trial." On each trial the organism may do one of several things. If it is in state 1, it may remain in state 1 (by pecking or sitting still in the example of the pigeon) or it may go to state 2. If the organism is in state 2, it may either remain there or go to state 1. These possibilities are shown in Fig. 4.

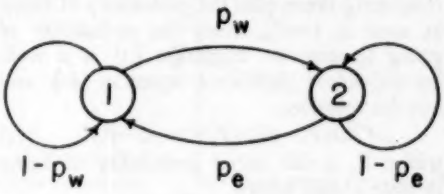

FIG. 4. DIAGRAM OF THE RESPONSE MODEL DISCUSSED

States 1 and 2 are denoted by the circles and the arrows show the possible transitions. The probability of going from state 1 to state 2 is p_w and the probability of going from state 2 to state 1 is p_e.

The probabilities of the various acts are shown in the figure; p_w is the probability of going from state 1 to state 2 (withdrawal from S) when the organism is in state 1, and p_e is the probability of going from state 2 to state 1 (entering S) when it is in state 2. If the probabilities p_w and p_e are constant, we have set up the problem of a simple two-state Markov chain (8). We want to compute the absolute probability of finding the organism

in state 1 at the start of trial n when $n=0, 1, 2, \ldots$. This will allow us to estimate the proportion of trials in which the organism is in situation S (state 1). This problem has been solved frequently in the mathematical literature in recent years. We will only sketch the proof here.

We let P_n be the probability of being in state 1 at the start of trial n. This probability on the next trial is given by:

$$P_{n+1}=P_n(1-p_w)+(1-P_n)p_e. \qquad [3]$$

In words, this equation says that the probability of being in state 1 at the start of the $(n+1)$st trial is equal to that probability, P_n, for the n-th trial times the probability of remaining there, plus the probability of being in state 2, $1-P_n$, times the probability of going to state 1. Equation [3] is a well-known linear difference equation (10) and has the solution:

$$P_n=P_\infty-(P_\infty-P_0)(1-p_w-p_e)^n \qquad [4]$$

where P_0 is the initial probability of being in state 1, and where

$$P_\infty=\frac{p_e}{p_e+p_w} \qquad [5]$$

is the asymptotic probability of being in state 1. We shall assume that equilibrium has been established so that the probability of finding our organism in state 1 is given by equation [5].

We now let ω_0 be the number of trials per minute, i.e., $\omega_0=1/h$ while h is the duration of a trial in minutes. Hence, by equation [5] the average number of trials per minute which begin in state 1 (situation S) is

$$\omega=\frac{p_e}{p_e+p_w}\omega_0. \qquad [6]$$

The probability of response r in state 1 is p_r and we must have

$$p_r+p_w\leq 1. \qquad [7]$$

Finally, we denote the number of r responses per minute by f and this quantity is

$$f=\omega p_r=\frac{p_e p_r}{p_e+p_w}\omega_0. \qquad [8]$$

Equation [8] plays the role in our model of Miller's assumption that the net response strength is the strength of that response minus the strength of conflicting responses.

In the pigeon experiment of Brush, *et al.*(1) ω_0 in equation [8] is the rate of pecking (pecks per minute) after complete conditioning but before punishment and f is the rate of pecking when the bird is in conflict. The three probabilities p_e, p_r, and p_w depend upon the similarity index, η, and these dependence must be determined from the set-theoretic model described above. We have already discussed the range of functional dependence of p_r and p_w on η. There remains, then, the need for determining—or postulating—how p_e depends on η. Once this is accomplished we will be in a position to predict the range of variations of the frequency f of pecking as a function of η.

It will be recalled that p_e is the probability that the organism will enter state 1 when it is in state 2, e.g., the probability that our pigeon will go to the immediate neighborhood of the window when the bird is in the back of the cage. Thus, we should expect p_e to depend upon whether or not entering state 1 has been reinforced and if so, how much. In other words, the value of p_e will depend upon the particular conditions of reinforcement and punishment used in the experiment. Our main interest, however, is in how p_e depends upon the similarity index, η. How does p_e change when S is changed to some S'? To answer this question, we make a tentative assumption: when the stimulus changes are restricted to the situation S, i.e., when no stimulus changes occur in state 2, then, p_e is unchanged, is independent of η. When the conditions of this hypothesis are not met, p_e may depend upon η, but we shall not consider this possibility in what follows. With the assumption that p_e is independent of η we may let $K=1/p_e=$ constant, and write equation [8] in the form:

$$\frac{f}{\omega_0}=\frac{p_r}{1+Kp_w}. \qquad [9]$$

We shall consider K a parameter in our model, but we must have from its definition, $K\geq 1$.

We now consider the two limiting cases of Fig. 3 for the variation of p_r and p_w with η. We let $K=2$ ($p_e=1/2$) and take $p_r=p_w=0.5$ in S (at $\eta=1$). The limiting cases of Fig. 3 then give for the ratio f/ω_0 the curves shown in Fig. 5(a) and 5(b). For the intermediate case when p_r and p_w are both linear with η, as in Fig. 2(c), we obtain the curve shown in Fig. 5(c). The functional dependence indicated in Fig. 5(b) could be called, in an operational sense, displacement. The conditions necessary for its appearance are analogous to those assumed by Miller (12): the probability of withdrawal must decrease

more rapidly than the probability of response r as similarity is decreased.

The position and height of the maximum in Fig. 5(b) depends of course upon the functions p_r (η) and p_w (η) as well as upon the parameter K. We consider now only the limiting case analogous to Fig. 3(b) for which p'_w decreases linearly with η until it becomes zero at the point $\eta=p_w$ (value at $\eta=1$) and for which p'_r is constant for $1\geq\eta\geq p_w$. This is the case which gives the greatest possible degree of displacement. It is easy to show for this case that the largest value of the ratio f/ω_0 is:

$$\left(\frac{f}{\omega_0}\right)_{max}=p_r \qquad [10]$$

and that this maximum occurs at $\eta=p_r$.

In the experiments reported by Brush, *et al.* (1), the stimulus situation was altered by changing the size of the light spot on the window pecked by the pigeon. The cues in the vicinity of the window were very likely conditioned mainly to pecking rather than withdrawal since shock was administered through the bird's feet. Therefore, it seems reasonable to us to suppose that the state of affairs depicted in Fig. 2(a) occurred: p_r decreased more rapidly than p_w. If this were the case our model predicts a monotonic decrease in response frequency with decreasing similarity. Indeed, this is what Brush, *et al.* observed.

In a more recent set of experiments, Jenkins, John, and Watson [2] repeated the procedure used by Brush, *et al.*, except that the floor texture and color were changed instead of the size of the light spot. Their hypothesis was that the cues which were changed were more strongly conditioned to withdrawal than to pecking, and hence the authors created what they considered to be the best conditions for observing a displacement effect. If that view were correct, Fig. 2(b) would represent the approximate variations of the probabilities and a variation of response frequency nearer to that in Fig. 5(b) should have obtained. Jenkins, John, and Watson did not observe significant increases in pecking rate when the floor of the bird's cage was changed, but the data did show a tendency more like that shown in Fig. 5(b) than did the data from the Brush, *et al.* study.

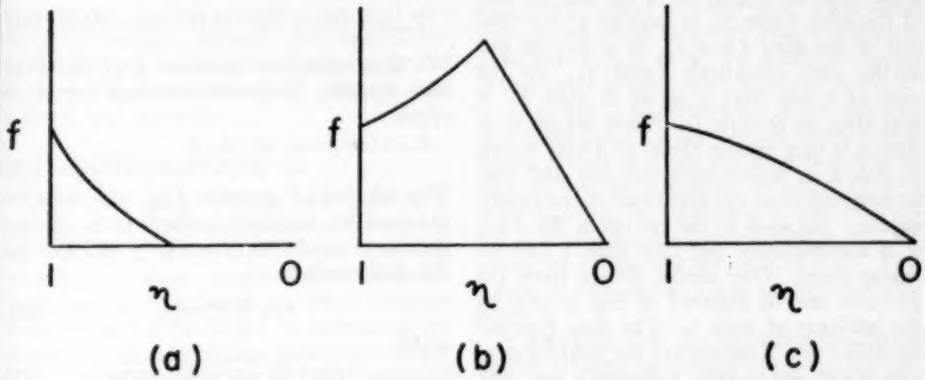

FIG. 5. FREQUENCY OF RESPONSE r (IN ARBITRARY UNITS OF ω_0) AS A FUNCTION OF THE SIMILARITY INDEX FOR THE CASES SHOWN IN FIG. 3(a), 3(b), and 2(c)

The Miller-Kraeling Experiment

Miller and Kraeling (13) reported an experiment of a somewhat different type in order to test Miller's displacement hypothesis. Rats were conditioned to run down an alley and push open a small door to get food. After this training, the rats were shocked through the door when the door was opened, for several trials until very few of the rats used opened the door in the allotted time. Then, the rats were "tested" with alleys of different widths and colors. With a particular test alley, four trials of 60-sec. duration were used. The data reported were proportion of rats in the population which opened the door on one or more of the four test trials.

[2] JENKINS, W. O., JOHN, W. F., & WATSON, P. Paper in preparation.

Miller and Kraeling found that that proportion with the "different" alley was significantly greater than with the "same" alley. This was taken as evidence of displacement.

As was pointed out in the Brush, *et al.* paper (1) the measure of response strength used by Miller and Kraeling is not directly comparable to the measure used in the pigeon studies discussed above. For this reason we must now extend our mathematical model in order to handle the Miller-Kraeling data. We conceive that there are three "states" for a rat to be in: it may be in the starting end of the alley (state 2), it may be at the door end of the alley (state 1), or it may be out of the alley completely (state 3). At the start of a test trial, a rat is in state 2. It may then go to state 1. When the rat is in state 1, it may remain there, or it may return to state 2, or it may open the door and then be removed from the apparatus by the experimenter. As soon as the rat opens the door, it is automatically put into state 3 and remains there. Our model differs from the previous one, as depicted in Fig. 4, only by the addition of state 3. The data reported by Miller and Kraeling are the proportion of rats which get to state 3 during a test trial.

We must now distinguish between a "test trial," as the term is used in the last paragraph, and the "trial of length h" which entered our earlier analysis. A "trial of length h" is an interval of time just long enough for a rat to go from one state to another—of the order of magnitude of one second in the Miller-Kraeling study. Hence, a "test trial" of 60-sec. duration constitutes roughly 60 "trials of length h." In the following discussion, we will use the word "trial" to denote a "trial of length h."

The presence of the absorbing state in our model (state 3) complicates the mathematical analysis a bit. As before we let P_n be the proportion of animals in state 1 at the start of trial n. We further define T_n to be the proportion of animals in state 3 at the start of trial n. Hence, the proportion of animals in state 2 at the start of trial n is $(1-P_n-T_n)$. The probability of going from state 2 to state 1 is p_e, the probability of going from state 1 to state 2 is p_w, and the probability of going from state 1 to state 3 is p. Other transitions, from 2 to 3, from 3 to 1, and from 3 to 2 are not allowed, i.e., have zero probability. With these definitions, we can write down the two equations,

$$T_{n+1}=T_n+pP_n, \qquad [11]$$

$$P_{n+1}=P_n(1-p-p_w)+(1-P_n-T_n)p_e. \qquad [12]$$

The solution of this pair of difference equations is straightforward but a bit tedious. The procedure is as follows: we solve equation [12] for T_n to obtain:

$$T_n=\frac{1}{p_e}[(1-p-p_w-p_e)P_n+p_eP_{n-1}]. \qquad [13]$$

From this result we then write:

$$T_{n+1}=\frac{1}{p_e}[(1-p-p_w-p_e)P_{n+1}+p_eP_n]. \qquad [14]$$

We then substitute equations [13] and [14] into equation [11] and rearrange terms to obtain:

$$P_{n+2}-(2-p-p_w-p_e)P_{n+1}+(1-p-p_w-p_e+pp_e)P_n=0. \qquad [15]$$

The solution of equation [15] may now be obtained by standard methods (10). Using the initial conditions that $P_0=0$ and $P_1=p_e$, the final result is:

$$P_n=\frac{p_e}{a}(r_1^n-r_2^n), \qquad [16]$$

where

$$a=\sqrt{(p+p_e+p_w)^2-4pp_e}, \qquad [17]$$

$$r_1=\frac{2-p-p_e-p_w+a}{2} \qquad [18]$$

$$r_2=\frac{2-p-p_e-p_w-a}{2}. \qquad [19]$$

When equation [16] is then inserted into equation [11] and the resulting difference equation solved, we obtain finally:

$$T_n=\frac{pp_e}{a}\left[\frac{1-r_1^n}{1-r_1}-\frac{1-r_2^n}{1-r_2}\right] \qquad [20]$$

where a, r_1, and r_2 are defined above. Our final result, equation [20] gives us an expression for the proportion of organisms to be found in state 3 after n trials. Of course, to compute T_n, we must know the three transition probabilities, p, p_e, and p_w. Our main interest is in how this quantity T_n depends on the similarity index η.

A numerical example will illustrate the dependence of T_n on η. Assume that in the "same" alley we have conditioned and punished a rat such that $p=0.01$, $p_w=0.49$, and $p_e=0.50$. For $n=60$, equation [20] yields $T_{60}=0.26$. Now assume that the "different" alley has the effect of reducing p_w to

zero but leaves $p=0.01$ and $p_\omega=0.50$. With these values and with $n=60$, equation [20] gives $T_{60}=0.43$. Therefore, the proportion of animals which open the door in 60 trials (one "test trial") is greater in the "different" alley than in the "same" alley. This conclusion is consistent with the results obtained by Miller and Kraeling. Note that we have not attempted to fit their data by estimating the parameters in our model from the data. Rather, we have given a numerical example which shows that our model can lead to numbers such as those observed. The crucial assumption in that example was that the stimulus change reduced the withdrawal probability p_w, but did not change the probability p, that the animal would open the door when it was in state 1.

A Cross-Cultural Problem

In a recent study Whiting and Child have attempted to apply displacement theories to the analysis of cross-cultural data (16). The societies of their sample were classified according to the degree to which diseases were believed to be caused by various agents among which were ghosts (spirits in human form) and animal spirits. It was assumed by them that such beliefs might represent the displacement (and/or projection) to these supernatural beings of aggression from living people who had frustrated them. Furthermore they assumed that these supernatural beings could be placed on a similarity dimension in such a way that animal spirits were less similar to living frustrators than were ghosts.

The hypothesis put forth was that societies in which there was high anxiety about aggression would displace this aggression to objects quite dissimilar to people (animal spirits), while societies with low anxiety about aggression would displace it to objects more similar to people (ghosts).

To test the hypothesis societies were rated on the level of anxiety resulting from the severity of punishment of aggression during childhood and were then divided into high and low groups on this variable. In addition, the "custom potentials" of the belief that ghosts and animal spirits cause illness were computed from ratings; these were rough indices of the frequencies of verbal responses such as "illness is caused by ghosts" or "illness is caused by animal spirits."

The data obtained confirmed the hypothesis. Societies rated as having high anxiety more often ascribed diseases to animal spirits, while societies with low anxiety more often ascribed illness to ghosts. In addition to the predicted effect, the data showed that the "custom potential" for animal spirits was significantly higher for societies with high anxiety as compared to societies with low anxiety. A similar result was obtained by Wright (17) in a cross-cultural study of the aggressive content of folk tales. This effect was not predicted either by Miller's displacement theory (12) or by the model presented in this paper so far.[3] The "response strength" or frequency, f, of occurrence, of the response at a given point on the similarity continuum would always be lower for high punishment than for low punishment. This may be seen from equation [9] for example—the value of f/ω_0 is always less than or equal to p_r and the greater p_w the smaller the value of f/ω_0. This is true at any chosen value of the similarity index, η.

It appeared that the displacement model described above had failed to predict the differences found in the cross-cultural studies. The difficulty, we believe, is that the model employed a too restricted experimental paradigm to apply to the cross-cultural data. It was always assumed that an organism was trained in a situation S and was then "tested" in one other *single* situation, S', at a time. This paradigm fitted the experimental situations used by Brush, *et al.* (1) and by Miller and Kraeling (13), but it does not seem to be adequate for a description of displacement in a social environment. When a response towards a given object is inhibited in a social situation, the person is not "tested" in a *single* situation S' but is exposed to a variety of social objects having various degrees of similarity to the original object. In order to describe displacement in a social environment, we will generalize our model as follows. An organism is trained (rewarded and/or punished) in S and is then presented

[3] This effect was predicted, however, by Whiting and Child from an assumption concerning the effects of conflict on motivation. We will here suggest a different derivation which is not incompatible with that hypothesis but is more relevant to the kind of problems we are dealing with in this paper.

with k situations having various degrees of similarity to S. The organism may then choose among these various situations and may or may not make response r in one chosen. We let there be a neutral state, labeled state o, in which response r is not possible. From state o, the organism may enter a number of states, labeled states 1, 2, . . . , k. State 1 will correspond to situation S, the one in which the original training occurred. States 2, 3, . . ., k will correspond to S' situations and each will have a different similarity index η with respect to S. Transitions between states 1, 2, . . ., k are forbidden but the organism enters state i (i=1, 2, . . ., k) from state o with probability $p_o^{(i)}$. The probability of withdrawal from state i to state o is $p_w^{(i)}$.

The proportion of organisms in state i at the start of trial n is $P_n^{(i)}$ and the proportion in state o is $P_n^{(o)}$. We may write at once that

$$P_{n+1}^{(i)}=P_n^{(i)}(1-p_w^{(i)})+P_n^{(o)}p_o^{(i)}. \qquad [21]$$

We have the conditions that

$$\sum_{i=1}^{k} P_n^{(i)}+P_n^{(o)}=1, \qquad [22]$$

and

$$\sum_{i=1}^{k} p_o^{(i)}\leqq 1. \qquad [23]$$

We shall be concerned only with the equilibrium conditions and so we take $P_n^{(i)}=P_{n+1}^{(i)}=P_\infty^{(i)}=P^{(i)}$. We then have from equation [21],

$$P^{(i)}=P^{(o)}\frac{p_o^{(i)}}{p_w^{(i)}}. \qquad [24]$$

Next, we let $p_r^{(i)}$ be the probability of response r in state i and f_i be the frequency of response r in state i. We then have

$$f_i=\omega_o p_r^{(i)}P^{(i)}=\omega_o P^{(o)}\frac{p_r^{(i)}p_o^{(i)}}{p_w^{(i)}}, \qquad [25]$$

where ω_o is the frequency of response r in S after complete conditioning but prior to punishment. From equations [22] and [24] we also have

$$P^{(o)}=\frac{1}{1+\sum_{i=1}^{k}(p_o^{(i)}/p_w^{(i)})} \qquad [26]$$

and so equation [25] becomes

$$f_i=\frac{\omega_o}{1+\sum_{i=1}^{k}(p_o^{(i)}/p_w^{(i)})}\frac{p_r^{(i)}p_o^{(i)}}{p_w^{(i)}}. \qquad [27]$$

For k=1, this result is identical to equation [8] above.

As an example of the use of equation [27] consider the case of k=3, i.e., when the organism can choose among the original stimulus situation (state 1) and two similar situations (states 2 and 3). First, we consider the case of no punishment in S and assume that $p_w^{(i)}$ is the same for states 1, 2, and 3, and has a value very small compared to the $p_o^{(i)}$ which are also assumed equal for the three states. Equation [27] then becomes approximately

$$f_i\cong\frac{\omega_o}{k}p_r^{(i)} \qquad \text{(no punishment)}. \qquad [28]$$

For the case being considered, k=3. As a numerical example we choose 1.00, 0.60, and 0.20 as values of η for states 1, 2, and 3, respectively. Accordingly, since the $p_w^{(i)}$ are taken to be very small, we take $p_r^{(1)}$=1.00, $p_r^{(2)}$=0.60, and $p_r^{(3)}$=0.20. Equation [28] then gives, f_1=0.333ω_o, f_2=0.200ω_o, and f_3=0.067ω_o. As would be expected the frequency of responding decreases monotonically with decreasing similarity. Now consider a case of relatively high punishment and assume that the withdrawal probabilities decrease more rapidly with η than do the response probabilities, $p_r^{(i)}$. As a numerical example, we choose the same values of η used in the previous example and let $p_r^{(1)}$=0.50, $p_r^{(2)}$=0.40, and $p_r^{(3)}$=0.18. We require that $p_r^{(i)}+p_w^{(i)}=\eta$ for i=1, 2, 3, and so we obtain $p_w^{(1)}$=0.50, $p_w^{(2)}$=0.20, and $p_w^{(3)}$=0.02. We further let $p_o^{(1)}=p_o^{(2)}=p_o^{(3)}$=0.20. These assumed numbers when inserted into equation [27] give f_1=0.016ω_o, f_2=0.032ω_o, and f_3=0.145ω_o. By comparing these results with those obtained for the no punishment example, we observe the general effect observed by Whiting and Child—f_3 is greater for high punishment than for low punishment. We have made no attempt to reproduce the numbers from the cross-cultural study; we have wished only to illustrate that the extended form of our model described in this section is capable of yielding frequencies of responding which are related to one another in the same general way as the "custom potentials" obtained by Whiting and Child.

Even with the extension given in this section, we do not consider our model to be fully adequate for describing displacement in a social environment. First of all, it does not

consider the effect, discussed by Whiting and Child, of conflict on the strength of drive. Secondly, as was pointed out by Miller (12) and by Brush, *et al.* (1), an important consideration in the application of displacement theories is the effect of discrimination learning. Miller's model and ours implicitly assume that the learning occurs only in the original stimulus situation S (state 1), and that responses made in the various similar situations have negligible effect on the probabilities. For restricted experimental designs, this assumption seems justified. In the Brush, *et al.* study, the data indicated no significant discrimination learning—the "test" periods were short compared to the time for original acquisition of the pecking response. In clinical and anthropological applications of displacement theory, however, it is likely that the effects of relearning can not be completely neglected. The difficulties in constructing a model for this process are appreciable. Not only is the process a complex stochastic one, but the effects of the responses in one similar situation will generalize to all other situations. Moreover, it is not at all clear to us how one would adequately describe the reward values of responding to similar objects. Detailed consideration of secondary reinforcement might also be involved. In general, most of the outstanding problems in social learning arise in the construction of such a model.

Recovery from Punishment—A Related Problem

A problem closely related to the ones we have discussed is that of recovery from punishment. Estes (7) found that if rats were first rewarded and then punished for bar pressing, a temporary increase in the rate of bar pressing occurred during extinction. These results were confirmed in a recent study by Edwards (6).

The model presented above can be applied to this problem if an additional assumption is made about the effects of an unreinforced, unpunished occurrence of response r (e.g., bar pressing). Bush and Mosteller assumed (2) an extinction operator E, defined by

$$Ep=(1-b)p \qquad [29]$$

Accordingly, we shall assume that an occurrence of response r will change p_r to $\beta p_r (0<\beta<1)$ and will change p_w to $\alpha p_w (0<\alpha<1)$. We further assume that p_o is unchanged by response r occurrences and that no other events change any of the probabilities. These assumptions define a non-Markovian process since the probabilities in Fig. 4 are no longer constant (8). We have not developed a mathematical treatment of this process but we have made numerical computations using the "Monte Carlo" method. The procedure was as follows: initial probabilities and parameters were chosen. We took $p_r=0.2$, $p_w=0.8$, $\beta=0.9$, $\alpha=0.6$, $p_o=0.5$. On the first trial the organism was assumed to be in state 1. A four-digit number was chosen from a table of random numbers. If that number was below

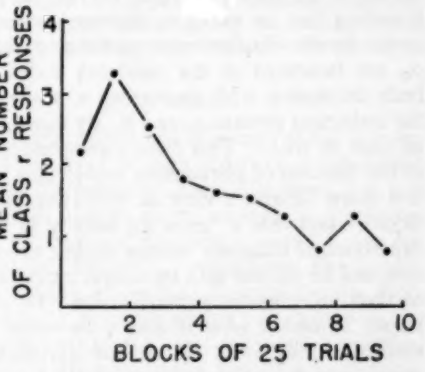

FIG. 6. MEAN NUMBER OF CLASS r RESPONSES IN BLOCKS OF TRIALS FOR 15 "MONTE CARLO" RUNS TO ILLUSTRATE THE THEORETICAL RECOVERY FROM PUNISHMENT

2000 then the hypothetical organism made response r and p_r was multiplied by 0.9 and p_w by 0.6, and the organism remained in state 1. If the random number was 2000 or greater, response w occurred, the probabilities did not change, and the organism went to state 2. On each trial a new random number (RN) was chosen. In state 1, if $0 \leq RN < 10{,}000 p_r$, then response r occurred; if $10{,}000 p_r \leq RN < 10{,}000(p_r+p_w)$, then response w occurred; if $10{,}000(p_r+p_w) \leq RN$, then response o occurred. Responses r and o left the organism in state 1 while response w took it to state 2. In state 2, if $RN<5000$ the organism went to state 1, while if $5000 \leq RN$, it remained in state 2.

The above procedure was repeated for 250 trials for each of 15 hypothetical organisms.

The number of occurrences of response r in blocks of 25 trials were counted for each run and the means for all runs are shown in Fig 6. The recovery effect is evident and its genesis is clear: during the early part of extinction, the organism spends more and more time in state 1 and this more than compensates for the decrease in p_r. Later in extinction, however, p_r continues to decrease and the frequency of r occurrences decreases accordingly.

From the point of view of our model this phenomenon of recovery from punishment is closely related to displacement. Both occur because p_w decreases rapidly causing the organism to be in state 1 more and more, while p_r decreases less rapidly, thereby producing a temporary increase in rate of responding but an eventual decrease towards zero. In the displacement problem, p_r and p_w are functions of the similarity index η, both decreasing with decreasing η, while in the extinction process, p_r and p_w are functions of time or trials. This close correspondence of the two sets of phenomena would also follow from Skinner's view of extinction (15). Skinner says that a "growing novelty in the experimental situation" occurs during extinction and he related data on simple extinction to data on stimulus generalization. In our terms, Skinner's view is that η decreases as extinction progresses. If we had introduced this as an assumption in our model, we could have derived punishment recovery from displacement or vice versa. Indeed, in our set-theoretic model, generalization and extinction are merely two ways of reducing the measure of elements conditioned to response r, and hence of reducing its probability of occurrence. If we had wished to do so, we could have described the stimulus situation after partial extinction as having a certain similarity η to the situation just prior to extinction.

SUMMARY

A model for displacement phenomena, suggested by Miller's formulation (12), is presented. The effects of stimulus generalization are described in terms of a set-theoretic model presented earlier (3). Punishment of a response, it is assumed, removes an organism from the situation and thereby reduces its opportunity for responding. From these assumptions, the conditions necessary for displacement are derived. The model is applied to recent experimental results on pigeons and rats (1, 13). The model is then extended to describe displacement in a social environment where an organism may respond in one of a number of stimulus situations. This extension of the model is compared with the cross-cultural study of Whiting and Child (16). Finally, the displacement model is related to the recovery from punishment effect observed by Estes (7).

REFERENCES

1. BRUSH, F. R., BUSH, R. R., JENKINS, W. O., JOHN, W. F., & WHITING, J. W. M. Stimulus generalizations after extinction and punishment: an experimental study of displacement. *J. abnorm. soc. Psychol.*, 1952, 47, 633–640.
2. BUSH, R. R., & MOSTELLER, F. A mathematical model for simple learning. *Psychol. Rev.*, 1951, 58, 313–323.
3. BUSH, R. R., & MOSTELLER, F. A model for stimulus generalization and discrimination. *Psychol. Rev.*, 1951, 58, 413–423.
4. DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., & SEARS, R. R. *Frustration and aggression.* New Haven: Yale Univer. Press, 1939.
5. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy.* New York: McGraw-Hill, 1950.
6. EDWARDS, R. P. The effects of punishment on unpunished responses. Ph.D. thesis, Harvard Univer., 1951.
7. ESTES, W. K. An experimental study of punishment. *Psychol. Monogr.*, 1944, 57, No. 3 (Whole No. 263).
8. FELLER, W. *Probability Theory and its Applications.* New York: Wiley, 1950. Pp. 307–346.
9. FREUD, S. *Basic writings.* New York: Modern Library, 1938.
10. JORDAN, C. *Calculus of finite differences.* New York: Chelsea, 1950.
11. MILLER, N. E. Experimental studies of conflict. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald, 1944. Pp. 431–465.
12. MILLER, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, 43, 155–178.
13. MILLER, N. E., & KRAELING, D. Displacement: greater generalization of approach than avoidance in a generalized approach-avoidance conflict. Paper read at East. Psychol. Ass., Brooklyn, March, 1951.
14. SEARS, R. R. Survey of objective studies of psychoanalytic concepts. *Soc. Sci. Res. Counc. Bull.*, 1943, No. 51.
15. SKINNER, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, 57, 193–216.
16. WHITING, J. W. M., & CHILD, I. L. *Child training and personality.* New Haven: Yale Univer. Press, 1953.
17. WRIGHT, G. C. Projection and displacement: a cross-cultural study of the expression of aggression in folktales. Ed.D. thesis, Harvard Univer., 1952.

Received May 29, 1952.

The Journal of Abnormal and Social Psychology
Vol. 48, No. 2, 1953

SMALL GROUP DISCUSSIONS WITH PARTICIPATORY AND SUPERVISORY LEADERSHIP[1]

A. PAUL HARE
Department of Sociology and Anthropology, Wellesley College

THE purpose of this research was to repeat with a different age-group an experiment by Preston and Heintz (2) which showed that participatory leadership was more effective than supervisory leadership as a technique in producing change of opinion in small discussion groups of college students.[2]

METHOD

A number of troops of Boy Scouts in a summer camp were told a story about a camping trip which would require each boy to travel alone through unknown country. They were then asked to rank 10 items of camping equipment in the order of their importance for such a trip. Next the troops were divided into groups of a leader and five followers for a 20-minute discussion period during which each leader recorded his group's decision concerning the importance of the items of camping equipment. Just before the discussions the leaders were taken aside and given instructions for participatory or supervisory leadership. After the discussions each individual ranked the equipment again and recorded his reactions to the discussion and to the leader on a questionnaire.

In addition to computing the rank-order correlations used by Preston and Heintz, within each group the average correlation of all rank orders, r_{av}, was computed before and after discussion (3, pp. 372–375).

The questionnaire was modified so that instead of reporting on influence, satisfaction, etc., the Scouts were asked to report good and bad things about the leadership and to describe what they would have done if they had been the leader.

Subjects

The subjects (*Ss*) were Boy Scouts who attended a Philadelphia Scout Camp during the summer of 1949. Since the Scouts usually came to camp with their home troops the members of all but one of the discussion groups knew each other the year round. The average age of the followers was 13 years. The leaders, boys from the same troop who had had some leadership experience, averaged 14 years.

A Nonrandom Factor in the Sample

For each group led by a supervisory leader there was not always another group from the same troop led by a participatory leader. As a result, 6 pairs of the groups in the sample were matched and 3 were unmatched; that is, they came from different troops. When the amount of change in opinion ($\bar{r}_{av}$ before$-\bar{r}_{av}$ after) was compared for the matched and unmatched groups a significant difference was found,[3] indicating a greater similarity in the amount of change in consensus when the supervisory and participatory groups come from the same scout troops than when they come from different troops.

RESULTS

A comparison of the means of the average correlation of all rank orders ($\bar{r}_{av}$) before and after discussion indicates that there is a significant change in the amount of agreement in each type of group as a result of the discussion (Table 1). Furthermore, the amount of change in the participatory groups (.54) is greater than the amount of change in the supervisory group (.37), although the difference is not significant, having a probability of $>.30$ of occurring by chance.

The resemblance between the initial and group ranking, $\bar{r}'_{12}$ (Table 2), is higher in every case than that reported by Preston and Heintz, suggesting that there is more initial agreement among Boy Scouts concerning camping equipment than among college students concerning the desirability of presidential candidates.

The average correlation of the group and final ranking, $\bar{r}'_{23}$, is also higher in every case than the corresponding correlation obtained by Preston and Heintz with the participatory followers showing significantly more agreement with the group ranking than the supervisory followers.

The participatory leaders have a significantly higher correlation for r'_{12} than do either the supervisory leaders or participatory

[1] Abstracted from HARE, A. P. A study of interaction and consensus in different sized discussion groups. Unpublished doctor's dissertation, University of Chicago, 1951. Pp. 144–168.

[2] The writer is indebted to Professor Preston for suggesting this research and for providing material and advice in the early stages of the experiment.

[3] In this experiment a result is considered significant if it has a probability of $<.10$ of occurring by chance.

TABLE 1

AVERAGE CORRELATION OF ALL RANK ORDERS FOR FOLLOWERS BEFORE AND AFTER DISCUSSION

LEADER TYPE	NUMBER OF GROUPS	$\bar{r}_{av}$ BEFORE	$\bar{r}_{av}$ AFTER	*Diff*	*t*	*p*
Supervisory	9	.28	.65	.37	4.18	<.01
Participatory	9	.27	.81	.54*	3.85	<.01

* This value for the difference between the $\bar{r}_{av}$ before and after discussion is similar to that obtained in an experiment by the *E* (1) in another camp using participatory groups of the same size.

followers, which, when combined with a high r'_{23} and r'_{13}, suggests that they influenced the group decision. This result, which is not obtained by Preston and Heintz, is, however, consistent with the way the discussions were conducted. Most of the Boy Scouts participating in this experiment were very interested in the subject for discussion. (In some cases the arguments touched off by the "camping game" continued for several days.) For this reason, the participatory leaders were generally anxious to give their views as well as to insure an equal chance for everyone else. On the other hand, the supervisory leaders who were specifically told to stay out of the discussion had little chance to influence the group. They tend to have the least agreement with the group ranking after discussion as well as a high degree of consistency bebetween their initial and final rankings.

The data in Table 3 supplement those given in Table 1 by providing another measure of the change in opinion which results from the two types of leadership. In every case the difference between the $\bar{r}'_{13}$ and $\bar{r}'_{23}$ is negative, indicating that the correlation between the initial and final rankings is smaller than the correlation between the group and final rankings. The difference is significant for all but the supervisory leaders.

TABLE 2

RANK CORRELATIONS AS A FUNCTION OF LEADERSHIP TECHNIQUES AND THE ROLE OF THE INDIVIDUAL

SUBJECTS	*N*	$\bar{r}'_{12}$	$\bar{r}'_{23}$	$\bar{r}'_{13}$
Leaders				
Supervisory	9	.50	.80	.72
Participatory	9	.77	.98	.78
Followers				
Supervisory	36	.55	.86	.61
Participatory	36	.56	.96	.60

TABLE 3

CORRELATION BETWEEN GROUP AND FINAL RANKINGS COMPARED WITH CORRELATION BETWEEN INITIAL AND FINAL RANKINGS

SUBJECTS	$\bar{r}'_{13}-\bar{r}'_{23}$*	*t*	*df*	*p*
Leaders				
Supervisory	−.20	−.52	16	>.50
Participatory	−.95	−3.93	16	<.01
Followers				
Supervisory	−.52	−3.93	70	<.01
Participatory	−.96	−6.55	70	<.01

* Differences are the equivalents of differences in *z* functions.

This suggests that each of the other three types of group members tends to be influenced more by the group ranking than by his initial ranking. Participatory leaders and followers show relatively more agreement with the group ranking than do the supervisory leaders and followers. Both the r_{av} and the r' data tend to substantiate the finding of Preston and Heintz that group activity under participatory leadership is more productive of change of opinion.

Questionnaire Responses

A summary of the responses to the questionnaire given after the final ranking indicates that participatory *S*s are generally better satisfied with the decision made by their group than are the supervisory *S*s. The differences between leaders and followers of each type, while not significant, are consistent with the findings of Preston and Heintz.

On the three 160-mm. linear rating scales all subgroups rate their group discussions as generally friendly and enjoyable, as interested in the task, and as efficient and productive. Although Preston and Heintz report statistically significant differences in the responses

in the direction of more enjoyable, interesting, and efficient discussion in the participatory groups, a series of *t* tests reveals no significant differences in the present experiment. The fact that differences are not reported may be due to a tendency for 13-year-old boys to withhold unfavorable comments about their peers from an adult experimenter. This same tendency toward a positive report is evident in the responses to open-end questions included to provide some descriptive data about the effects of the two types of leadership style. No matter what the leader style, the boys tried to say something good about it.

SUMMARY

When 13-year-old boys were used as *S*s in an experiment which compared the effects of supervisory and participatory leadership on group judgment in discussion groups of five members, the results generally substantiated the findings of a similar experiment using college-age *S*s. The data indicated that participatory leadership was more effective than supervisory leadership as a technique for changing opinion. In addition, the participatory leader generally had more influence on the group, a result not obtained with college *S*s. Although participatory *S*s were generally better satisfied with the results of the group decision than were the supervisory *S*s, the differences were not statistically significant.

REFERENCES

1. HARE, A. P. A study of interaction and consensus in different sized groups. *Amer. sociol. Rev.*, 1952, 17, 216–267.
2. PRESTON, M. G., & HEINTZ, R. K. Effects of participatory vs. supervisory leadership on group judgment. *J. abnorm. soc. Psychol.*, 1949, 44, 345–355.
3. WOODWORTH, R. S. *Experimental psychology*. New York: Henry Holt, 1938.

Received June 4, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

THE INFLUENCE OF INDIVIDUAL MEMBERS ON THE CHARACTERISTICS OF SMALL GROUPS [1]

WILLIAM HAYTHORN
Human Resources Research Center, Randolph Air Force Base

THERE has been increased emphasis in social psychology recently on the description and measurement of group characteristics. Hemphill and Westie (10) and Cattell and Wispé (6) have been particularly concerned with isolating dimensions along which groups vary. However, little has been done to relate these group characteristics to the behavior of group members.

Cattell (4) has suggested that leadership be defined in terms of the effect the individual has on group "syntality," and has further hypothesized that each member of a group contributes something to the characteristics of the group. Redl (14) has discussed leadership in terms of the "central person" around whom group formative processes occur, the implication being that the central person is the primary factor in determining the nature of the group.

Other writers (12, 13, 15) in the area of group dynamics have theorized that group characteristics grow out of social interaction in the group, but to the author's knowledge there have been no experimental studies specifically focused on the relationships between the behavior of individual group members and the characteristics of the group. The present study attempts to explore some of these relationships.

METHOD

In order to isolate the effects an individual has on groups, it is necessary to have him work in several groups with different co-workers. Our experimental design, presented in Table 1, permits isolating the effect each subject (*S*) had on his groups, since each *S* worked in five unique groups; that is, no other *S* worked in more than one of them.

The *Ss* were NROTC sophomores, all of whom volunteered and were paid one dollar per hour.

TABLE 1

EXPERIMENTAL DESIGN

SESSION	GROUPS *			
1	ABCD	EFGH	IJKL	MNOP
2	AEIM	BFJN	CGKO	DHLP
3	AFKP	BELO	CHIN	DGJM
4	AGLN	BHKM	CEJP	DFIO
5	AHJO	BGIP	CFLM	DEKN

* Letters represent *Ss*.

They were brought to the laboratory in groups of four and asked to work on three different tasks—reasoning (R), mechanical assembly (MA), and discussion (D). The R and MA tasks are similar to those described in a previous publication (2). That is, the R task involved group solution of syllogistic reasoning problems, and the MA task involved building various structures from precut and predrilled lumber. Both tasks required a high degree of group participation and coordination. For the D task, TAT pictures were presented and the *Ss* were asked to compose a story as a group (9, 11). This invariably touched off an active, lively discussion. Group agreement was required.

While the *Ss* were working on the tasks, two observers (*Os*) independently watched and recorded behavior by a new recording technique, reported by Carter, *et al.* (3), in which the observers "categorize" and "type out" behavior as it occurs. At the end of each task the *Os* rated the *Ss* on twelve behavioral characteristics such as cooperativeness, aggressiveness, and efficiency. Reliability of rating scales ranged from .10 (cooperativeness on the MA task) to .98 with an average reliability of .86.

At the end of the session, both *Ss* and *Os* rated the group on eleven group characteristics such as morale, productivity, and cohesiveness. The average interobserver reliability for these ratings was .70, with a range from .20 to .92. The average correlation between ratings by *Os* and by *Ss*, which provides another kind of reliability measure, was .64. Although these reliabilities are lower than one would wish, they are considered high enough to make the ratings of some value in an exploratory study such as this one.

In addition, the *Ss* were asked to select the best and poorest leaders in the group, as well as the person with whom they liked best to work and the person with whom they liked least to work. Finally, after all groups had been run, *Ss* were asked to fill out the Cattell Sixteen Personality Factor Questionnaire (5).

[1] The work described in this paper was done under a contract between the U. S. Navy, Office of Naval Research, and the University of Rochester. This is the eighth paper resulting from that work, and is a revision of a portion of a doctoral dissertation presented at the University of Rochester in 1952 (9). The work was under the general supervision of Dr. Launor F. Carter. Drs. Beatrice Shriver and John Lanzetta rendered invaluable service as observers and informal consultants.

There were, then, the following measures of the behavior of individual group members:

1. The number of choices and rejections received by each *S* from his 15 co-workers on the two choice criteria. Because these were highly intercorrelated, a single measure of choice value was computed for each *S* by simply combining the two measures.

2. The ratings by two independent *O*s of each *S* on 12 variables on each of 3 tasks.

3. The behavioral or interaction categories recorded by *O*s for each *S*.

4. Scores on each of the 16 personality factors of the Cattell questionnaire.

For measures of characteristics of groups in which each *S* worked, two sets of scores were computed:

1. The ratings of group characteristics by each of the *S*'s 15 co-workers were summed. Thus, from Table 1 it can be seen that the average cohesiveness of groups in which subject A worked can be determined by summing the ratings of group cohesiveness by *S*s B_1, C_1, D_1, E_2, I_2, M_2, F_3, K_3, P_3, G_4, L_4, N_4, H_5, J_5, and O_5. The numeral subscripts indicate the session number.

2. The ratings of group characteristics by *O*s were summed over the five groups in which each *S* worked. The average morale, for example, of groups in which subject E worked can be determined by summing the ratings of group morale by the *O*s for groups EFGH, AEIM, BELO, CEJP, and DEKN.

TABLE 2

CORRELATIONS BETWEEN SUBJECTS' TOTAL SOCIOMETRIC SCORES AND RATINGS OF CHARACTERISTICS OF GROUPS IN WHICH SUBJECTS WORKED

($N=16$)

GROUP CHARACTERISTICS RATED	SOURCE OF RATINGS	
	CO-WORKERS	OBSERVERS
1. Talkativeness	−.24	.03
2. Morale	.68	.56
3. Competitiveness	−.10	−.18
4. Cooperativeness	.69	.38
5. Productivity	.36	.29
6. Cohesiveness	.52	.19
7. Motivation	.67	.27
8. Friendliness	.31	.09
9. Activity	.24	.14
10. Job Completion	.59	.49
11. Social Interaction	.30	.11

$r.05=.49$; $r.10=.43$.

RESULTS

Relationships between Ss' "Choice Values" and Characteristics of Groups in Which They Worked

Each *S*'s choice value was determined by counting all of the choices received by him from his co-workers, and subtracting from that all the rejections he received. Correlation coefficients between these scores and the average ratings of characteristics of groups in which each *S* worked (described above) were computed, and are presented in Table 2. The correlations with co-workers' ratings of group characteristics indicate statistically significant relationships between the degree to which an *S* was chosen by his co-workers and the extent to which co-workers rated groups in which he worked as high in morale, cooperativeness, cohesiveness, motivation, and interest in job completion. These results suggest that individuals who are highly chosen by co-workers "facilitate" group functioning, and that individuals who "depress" group functioning are not generally chosen by other members of the group. A possible alternative interpretation is that, since co-workers' ratings of group characteristics are correlated with co-workers' choices of group members, the results simply reflect halo in the co-workers' ratings. The former interpretation, however, is supported by the statistically significant correlations between *S*s' choice-value scores and ratings by *O*s of group morale and interest in job completion.

Relationships between Rated Behavior of Ss and Rated Characteristics of Groups in Which Ss Worked

The preceding results suggest that highly chosen persons contribute to characteristics indicative of "smooth" group functioning. In order to determine what kinds of *behavior* facilitate or depress each of the group characteristics rated, correlations were computed between the average ratings of each *S* by the *O*s on each of the 12 behavioral traits and the ratings by co-workers of the characteristics of groups in which the *S*s worked. The correlations, computed separately for the three tasks, are given in Table 3.

How the standard error of r for Table 3 should be determined is difficult to ascertain. There was an N of only 16 *S*s, but each *S* was observed in five different groups. An N of 80 gives an underestimation of standard error, since there were not 80 independent measures. An N of 16 gives an equally inaccurate overestimate of standard error, since there were 16 clusters of five scores each in the computations. However, since the latter

TABLE 3

CORRELATIONS BETWEEN RATINGS BY OBSERVERS OF BEHAVIORAL TRAITS OF SUBJECTS AND RATINGS BY SUBJECTS' CO-WORKERS OF CHARACTERISTICS OF GROUPS IN WHICH THE SUBJECTS WORKED

BEHAVIORAL TRAITS RATED	GROUP CHARACTERISTICS RATED										
	Talkativeness	Morale	Competitiveness	Cooperativeness	Productivity	Cohesiveness	Motivation	Friendliness	Activity	Job Completion	Social Interaction
Cooperativeness											
R	—	—	—	—	—	—	—	—	—	—	—
MA	—	—	—	—	—	—	—	—	—	—	—
D	−.30	.66	−.26	.47	.33	.30	.33	−.02	.29	.27	.27
Aggressiveness											
R	−.20	−.04	.00	−.12	−.05	−.21	−.12	−.27	.26	−.14	.05
MA	−.30	.33	−.24	.25	.18	−.04	.22	−.31	.28	.17	.16
D	.06	.18	−.35	−.05	.10	−.24	−.11	−.28	−.05	.18	.25
Initiative											
R	−.33	−.03	.01	−.11	.06	−.19	−.10	−.27	.32	−.03	.06
MA	−.37	.40	−.33	.44	.19	.17	.23	−.20	.31	.23	.32
D	−.08	.37	−.29	.16	.21	−.04	.15	−.23	.09	.00	.14
Prestige											
R	−.54	.31	−.21	.42	.31	.16	.31	−.11	.21	.39	.42
MA	−.47	.45	−.18	.47	.37	.22	.45	−.20	.30	.43	.27
D	−.19	.39	−.29	.41	.32	.13	.35	−.23	.10	.27	.29
Confidence											
R	−.33	−.05	.21	.04	.03	−.07	−.02	−.27	.10	−.04	−.15
MA	−.37	.30	−.27	.28	.10	.03	.24	−.32	.26	.15	.13
D	.12	.22	−.21	.10	.29	−.13	.10	−.23	−.09	−.12	.23
Submissiveness											
R	.31	−.01	−.05	.03	.07	.10	.08	.23	−.21	.10	.11
MA	.28	−.30	.32	−.33	−.14	−.05	−.29	.23	−.21	−.16	−.16
D	.00	−.14	.44	−.13	−.09	.10	.02	.28	.19	.14	−.29

$N=16$ $r_{.05}=.49$
$r_{.10}=.42$ $r_{.01}=.61$

TABLE 3—*Continued*

Behavioral Traits Rated	Group Characteristics Rated										
	Talkativeness	Morale	Competitiveness	Cooperativeness	Productivity	Cohesiveness	Motivation	Friendliness	Activity	Job Completion	Social Interaction
Efficiency											
R	−.49	.33	−.45	.48	.26	.27	.23	.18	−.04	.28	.49
MA	−.58	.52	−.33	.54	.37	.35	.34	−.10	.34	.41	.26
D	−.38	.47	−.11	.45	.20	.37	.54	−.10	.05	.35	.14
Sociability											
R	.36	.19	−.19	.00	.17	−.25	−.23	.11	−.11	−.11	.22
MA	.38	.14	.08	.01	.09	−.24	−.06	.24	−.16	−.09	.27
D	.43	.31	−.39	.15	.17	−.16	−.23	.16	.05	.13	.43
Insight											
R	−.54	.21	−.29	.28	.34	.10	.12	−.33	.09	.26	.21
MA	−.54	.48	−.22	.53	.35	.29	.48	−.14	.31	.45	.20
D	−.26	.45	−.28	.32	.31	.10	.31	−.20	.18	.16	.17
Authoritarianism											
R	−.27	−.05	.03	−.05	−.03	−.14	−.06	−.35	.26	−.09	.01
MA	−.12	.25	−.24	.20	.23	−.10	.19	−.34	.20	.11	.08
D	.00	.14	−.26	.12	.13	−.12	.08	−.35	−.09	−.04	.21
Interest in Individual Solution											
R	−.16	−.13	.14	−.16	−.11	−.22	−.08	−.30	.28	−.21	−.06
MA	−.24	.07	−.12	.03	−.11	−.18	.09	−.56	.35	−.01	−.05
D	.07	.09	−.04	−.08	.02	−.27	.44	−.27	.01	−.26	−.01
Leadership											
R	−.46	.02	−.01	.05	.12	−.12	.10	−.30	.28	.11	.13
MA	−.34	.34	−.14	.31	.21	.04	.34	−.31	.33	.24	.16
D	−.03	.47	−.40	.32	.37	.05	.28	−.10	.00	.15	.21

$N=16$ $r_{.05}=.49$
$r_{.10}=.42$ $r_{.01}=.61$

gives a more conservative estimate, it was adopted.

With an *N* of 16 it is obvious that not many of the relationships are statistically significant. However, there are many factors depressing these correlations. For one thing, both sets of ratings include considerable error variance, and no correction for attenuation has been applied. Secondly, the correlations are between ratings of limited aspects of the *Ss*' behavior and ratings of equally limited aspects of the group's characteristics. Presumably the *S*'s total behavioral pattern interacts with the total group in a more complicated manner than that indicated by these correlations. In the third place, the correlations indicate relationships between the rated behavior of single *Ss* and the rated characteristics of groups in which the *Ss* worked. It is more probable that the nature of a group is related to the behavior of all members and to the interaction between members. In view of this, and of the exploratory nature of the study, examining the obtained correlations for any suggested trends seems justified.

To discuss each of the correlations in Table 3 would be a tedious task, and would involve much repetition since the behavioral ratings are not completely independent. Carter and Couch (1) have factor analyzed ratings of similar traits rated under similar laboratory conditions, and have identified three factors which have been labelled Group Goal Facilitation, I; Striving for Individual Prominence, II; and Group Sociability, III. These factor analyses have been done with data from several experiments, using various *Os* and groups of various sizes and structures. There is, therefore, considerable reason to believe that most of the variance in the present 12 traits can be accounted for by the same three factors. Because of this, it is desirable to summarize the present results in terms of the factors.

If attention is directed to only those correlations significant at the .10 level or better, a definite pattern is evident. Of the 35 statistically significant correlations in Table 3, 27 are between Factor I traits—cooperativeness, prestige, efficiency, and insight—and group characteristics indicating a smooth group functioning. Group morale, for instance, was correlated .66 with individual cooperativeness on the D task, .45 with prestige on the MA task, .52 with efficiency on the MA task, .47 with efficiency on the D task, .48 with insightfulness on the MA task, and .45 with insightfulness on the D task. The only other trait significantly related to group morale was leadership (.47 on the D task). This suggests that Factor I behavior as seen by *Os* "facilitated" group morale as seen by co-workers.

Similarly, co-workers' ratings of group cooperativeness correlated .47 with individual cooperativeness on the D task, .42 with prestige on the R task, .47 with prestige on the MA task, .48 with efficiency on the R task, .54 with efficiency on the MA task, .45 with efficiency on the D task, and .53 with insightfulness on the MA task. Factor I behavior, then, apparently facilitated group cooperativeness.

Group motivation correlated .45 with ratings of prestige on the MA task, .54 with efficiency on the D task, and .48 with insightfulness on the MA task. Group interest in job completion correlated .43 with prestige on the MA task, and .45 with insightfulness on the MA task. These findings suggest that individuals whom *Os* rate as efficient, insightful, and having prestige also increase the motivation of other members of the group.

Group talkativeness, which *Ss* apparently considered inefficient activity, was negatively related to individual prestige on the R task (−.54) and on the MA task (−.47); to individual efficiency on the R task (−.49), and on the MA task (−.58); and to individual insightfulness on the R task (−.54) and on the MA task (−.54). Factor I behavior on the R and MA tasks significantly "depressed" group talkativeness, as rated by *Ss*' co-workers.

Prestige and efficiency on the R task were also positively related to group interest in social interaction (.42 and .49 respectively). It would seem from all these *r*'s that when one member of a group engaged in a high degree of "group goal-facilitating" behavior, the entire group functioned better in terms of the measures obtained.

Group friendliness, however, was negatively related to most of the individual traits rated. Factor II traits—aggressiveness, initia-

tive, confidence, authoritarianism, interest in individual solution, and leadership—tended especially to depress group friendliness. This factor—striving for individual prominence—seems to be somewhat similar to the self-oriented need behavior described by Fouriezos, Hutt, and Guetzkow (7). The finding that such behavior is negatively related to group friendliness is consistent with their finding that such behavior was negatively related to participant satisfaction with the meeting. Particularly relevant here is the correlation of —.56 between interest in individual solution and group friendliness.

Relationships between Personality Factors and Group Characteristics

It will be recalled that *Ss* were given the Cattell Sixteen Personality Factor Questionnaire after all the groups had been run. Correlations were computed between scores on the 16 factors and ratings of group characteristics by *Os* (described above). These correlations are presented in Table 4.

Before any of the correlations were computed, predictions were made about the direc-

TABLE 4

CORRELATIONS BETWEEN SIXTEEN PERSONALITY FACTOR QUESTIONNAIRE AND RATINGS OF GROUP CHARACTERISTICS BY OBSERVERS

(N=16)

GROUP CHARACTERISTICS RATED	PERSONALITY FACTORS										
	A	C	E	F	G	H	L	M	O	Q_2	Q_3
1. Morale	.24	.58				.10			—.49	—.11	
2. Competition	—.45										
3. Productivity	.17	.48	*—.16*	—.10	.19			—.61		.14	.41
4. Cohesiveness	.37		*.26*	.07		.29	—.69		—.33	*.15*	
5. Motivation		.27		—.25					—.26		
6. Friendliness			*.01*	.18		.17	—.57			—.33	
7. Job Completion		.43	*—.21*	—.21				—.43		.00	
8. Social Interaction			*.33*			.62				—.42	
$r_{.05}$=.49											

* For descriptions of the personality factors, see the text and Cattell (5). Italics indicate cases in which the direction of the relationship was predicted incorrectly.

There was some tendency for Factor III—sociability—to increase group talkativeness and interest in social interaction but the trend is not very significant statistically.

In summary it can be said that Factor I behavior—group goal-facilitation—was positively related to co-workers' ratings of group morale, cooperativeness, motivation, and interest in job completion. Factor II behavior—striving for individual prominence—tended to be negatively related to group friendliness. Factor III behavior—sociability—tended to be positively related to group talkativeness and interest in social interaction. Other specific findings, such as the correlation of .44 between individual submissiveness and group competitiveness, may suggest possible research leads for the future, but standing alone as they do here does not inspire much confidence in them.

tions of the relationships. The predictions were based chiefly on the general hypothesis that more mature, flexible, accepting persons would facilitate effective group functioning. A chi-square test of the hypothesis that there was no relationship between the predicted and computed signs of the correlations yielded a probability value of less than .001. That is, there is very little chance that the number of correct predictions resulted from chance. The alternate hypothesis that the directions of the relationships were predicted more accurately than chance by the general hypothesis stated must be considered tenable. The correlations in Table 4 therefore deserve a closer investigation. Only those correlations for which predictions were made were computed. Cases in which the direction of

the relationship was incorrectly predicted are italicized. All of the 10 statistically significant correlations (.10 level) are in the expected directions.

Factor A, Cyclothymia vs. Schizothymia (Cattell's labelling), involves cooperativeness, trustfulness, and adaptability. The correlations are all in the expected directions, suggesting a tendency for groups in which A+ *Ss* worked to be rated as having high morale (r=.24), low competitiveness (r=−.45) and high cohesiveness (r=.37).

Factor C, Emotional Stability or Ego Strength vs. General Neuroticism, was expected to correlate positively with those group characteristics indicating efficient group functioning. The maturity and stability of high C *Ss* were expected to lend stability and efficiency to groups in which C+ persons worked. The correlations between factor C scores and *Os*' ratings of group morale (.57), productivity (.47), and interest in job completion (.43) support these expectations.

Factor E, Dominance or Ascendance vs. Submission, was expected to be related positively to productivity and interest in job completion but negatively to group cohesiveness, friendliness, and interest in social interaction. All five correlations came out contrary to expectations. No satisfactory explanation of this prediction error has been achieved. The *Ss* simply did not react to a dominant group member as the experimenter expected.

Factor F, Surgency vs. Desurgency, is said to describe persons "carefree, happy-go-lucky, fond of bustle and excitement, inclined to practical jokes and disinclined to occupations requiring close and accurate work" (5). It was predicted that scores on factor F would be positively related to ratings of group cohesiveness and friendliness, but negatively related to group productivity, motivation, and interest in job completion. The correlations are in the predicted directions, but low and insignificant.

Factor H, Adventurous Cyclothymia vs. Inherent Withdrawn Schizothymia, is described as a measure of the extent to which the respondent likes people. The expectation that factor H scores would be positively related to group cohesiveness and interest in social interaction was strongly supported by the correlations (.29 and .62 respectively).

Factor L, Paranoid Schizothymia vs. Trustful Accessibility, involves paranoid suspiciousness, jealousy, dourness, and rigidity. Persons scoring high on this factor were expected to depress group cohesiveness and friendliness. The correlations (−.69 and −.57 respectively) clearly tend to confirm this expectation.

Factor M, Bohemianism vs. Practical Concernedness, allegedly measures unconventionality, eccentricity and undependability. The M+ individual is described as one who "unconcernedly goes his own way in the community and does not feel much responsibility in practical matters" (5). The correlations with *Os*' ratings of group productivity (−.61) and interest in job completion (−.43) lend strong support to the hypothesis that "practical, logical, conscientious" persons, low on factor M, facilitate group productivity while M+ individuals depress it.

Factor O, Worrying Suspiciousness vs. Calm Trustfulness, attempts to measure depressive tendencies, moodiness, etc. The correlations between factor O scores and *Os*' ratings of group morale (−.49), cohesiveness (−.33), and motivation (−.26) suggest that O+ persons detract from the pleasant feeling tone of the group.

Factor Q_2, Independent Self-Sufficiency vs. Lack of Resolution, is described as a measure of independence and resoluteness. A Q_2 person "prefers to work and make decisions in company with other people, likes social approval and admiration," etc. The correlations between factor Q_2 scores and *Os*' ratings of group friendliness (−.33) and interest in social interaction (−.42) suggest that individuals who are highly self-sufficient and do not particularly care for social approval tend to depress the social aspects of their groups.

In describing factor Q_3, Will Control and Character Stability, Cattell writes that ". . . it has been observed that when people who are high in this factor have been put together in groups, the general efficiency and objectiveness of the group is decidedly higher than in groups that are only average in the factor" (5, p. 16). The correlation (.41) with *Os*' ratings of the productivity of groups in which Q_3+ persons worked lends support to the hypothesis that such persons facilitate group productivity.

In summary of the results in Table 4, it

seems probable that personality traits involving maturity, adaptability, and acceptance of others tend to facilitate or to be positively related to smooth and effective group functioning, while traits involving coolness, suspiciousness, eccentricity and the like tend to depress smooth group functioning (i.e., to be negatively related to ratings of group morale, cohesiveness, friendliness, etc.). This, of course, represents only a very general summary of the results, but it is felt that more specific conclusions cannot be drawn until much more research has been done. It seems clear, however, that *Ss*' responses to personality questionnaires are related to the effect *Ss* have on small groups, and that the characteristics of small groups are to some degree functions of the personality traits of individual group members.

DISCUSSION

The preceding results indicate that there are significant relationships between the behavior of individual group members and the characteristics of the total group, and that these relationships can be experimentally isolated. They also support Cattell's hypothesis that "every man in a group is to some extent a leader in so far as every man has some effect upon the syntality of a group" (4, p. 25). The present study suggests the possibility of studying leadership in terms of the relationships between individuals and group characteristics.

The results reported here are considered highly tentative. The lack of previous research in this area and the exploratory nature of this study greatly limit the extent to which the findings can be generalized. Within these limits, however, the following conclusions are considered tenable:

1. It is possible to isolate relationships between the behaviors of individual group members and the characterirtics or "syntality" traits of small groups.

2. Individuals who are chosen by coworkers as good leaders or as persons with whom others like to work "facilitate" group functioning, while individuals who "depress" group functioning are not generally chosen by other members of the group.

3. Individual behavior patterns which include cooperativeness, efficiency, and insight—Factor I behavior—tend to "facilitate" or be positively related to effective group functioning as measured by ratings of group morale, cooperativeness, productivity, cohesiveness, motivation, and interest in job completion.

4. Individual behavior patterns which include aggressiveness, self-confidence, initiative, interest in individual solution, and authoritarianism—Factor II behavior—tend to be somewhat negatively related to ratings of group cohesiveness and friendliness.

5. Sociable behavior—Factor III—tends to reduce group motivation and competition, but to increase group talkativeness, friendliness, and interest in social interaction.

6. Personality traits involving maturity, adaptability, and acceptance of others tend to be positively related to smooth and effective group functioning.

7. Personality traits involving suspiciousness, eccentricity, and coolness toward others tend to be negatively related to smooth group functioning.

The results of this study support Cattell's (4) and Gibb's (8) suggestions that leadership be studied as relationships between individual group members and "syntality" traits of the group. At least in small groups of this nature it seems probable that each individual member makes some contribution to the characteristics of the total group. Such contributions can be investigated with a design similar to the one used here. Experimental studies utilizing such a design need to be conducted with a larger number of *Ss*, different task situations, larger groups, and more refined measures of personality, behavior, and group characteristics. A knowledge of the effects of behavior of individuals on each other seems essential to an understanding of small group behavior.

SUMMARY

1. Sixteen male NROTC students met in groups of four. Each *S* worked with each other *S* once and only once, so that each *S* worked in five different groups. Each group worked on a Reasoning (R) task, a Mechanical Assembly (MA) task, and a Discussion (D) task.

2. The *Ss* were observed by two independent *Os* who rated individual behavior and group characteristics. The *Ss* also rated characteristics of their groups. Sociometric and personality data were obtained.

3. Relationships between measures of individuals and of groups were analyzed. Relationships were found, indicating that individual members significantly affect the characteristics of small groups. In general it was found that effective group functioning was facilitated by cooperativeness, efficiency, and insight, while behavior which we have called "striving for individual prominence" reduced group cohesiveness and friendliness.

4. Significant relationships in the predicted directions were also found between personality measures and ratings of group characteristics, suggesting that mature, accepting persons facilitate while suspicious, nonaccepting persons depress group characteristics indicative of smooth functioning.

5. Some of the implications of this study for research on leadership and small group behavior were briefly discussed.

REFERENCES

1. CARTER, L. F., & COUCH, A. Factorial studies of the rated behavior of group members. Paper read at East. Psychol. Ass., March 1952.
2. CARTER, L. F., HAYTHORN, W., & HOWELL, MARGARET. A further investigation of the criteria of leadership. *J. abnorm. soc. Psychol.*, 1950, **45**, 350–358.
3. CARTER, L. F., HAYTHORN, W., MEIROWITZ, BEATRICE, & LANZETTA, J. Note on a new technique of interaction recording. *J. abnorm. soc. Psychol.*, 1951, **46**, 258–260.
4. CATTELL, R. B. Determining syntality dimensions as a basis for morale and leadership measurement. In H. Guetzkow (Ed.), *Groups, leadership, and men*. Pittsburgh: Carnegie Press, 1951. Pp. 16–27.
5. CATTELL, R. B., SAUNDERS, D. R., & STICE, G. *Handbook for the sixteen personality factor questionnaire*. Champaign, Illinois: Institute for Personality and Ability Testing, 1949.
6. CATTELL, R. B., & WISPÉ, L. G. The dimensions of syntality in small groups. *J. soc. Psychol.*, 1948, **28**, 57–78.
7. FOURIEZOS, N. T., HUTT, M. L., & GUETZKOW, H. Measurement of self-oriented needs in discussion groups. *J. abnorm. soc. Psychol.*, 1950, **45**, 682–690.
8. GIBB, C. A. The research background of an interactional theory of leadership. *Aust. J. Psychol.*, 1950, **2**, 19–42.
9. HAYTHORN, W. The influence of the individual group members on the behavior of coworkers and on the characteristics of groups. Ph.D. thesis, University of Rochester, 1952.
10. HEMPHILL, J. K., & WESTIE, C. M. The measurement of group dimensions. *J. Psychol.*, 1950, **29**, 325–342.
11. HENRY, W. E., & GUETZKOW, H. *Group projection sketches for the study of small groups*. Publication No. 4 of the Conference Research Project. Ann Arbor: Univer. of Michigan, 1949.
12. NORTHWAY, M. L., FRANKEL, E. B., & POTASHIN, R. *Personality and sociometric status*. Sociometric Monographs, No. 11. New York: Beacon House, 1947.
13. POWELL, J. W. The dynamics of group formation. *Psychiatry*, 1948, **11**, 117–124.
14. REDL, F. Group emotion and leadership. *Psychiatry*, 1942, **5**, 573–596.
15. SLAVSON, S. R. *An introduction to group therapy*. New York: The Commonwealth Fund, 1943.

Received June 12, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

A PERSONALITY SCALE OF MANIFEST ANXIETY

JANET A. TAYLOR
Northwestern University

A SERIES of recent studies (3, 4, 5, 6, 7, 9, 10) has shown that performance in a number of experimental situations, ranging from simple conditioning and reaction time to a "therapy" situation involving experimentally induced stress, is related to the level of anxiety as revealed on a test of manifest anxiety. Most of these investigations were concerned with the role of drive or motivation in performance, drive level being varied by means of selection of subjects on the basis of extreme scores made on an anxiety scale rather than by experimental manipulation (e.g., electric shock, stress-producing instructions, etc.). The use of the anxiety scale in this connection was based on two assumptions: first, that variation in drive level of the individual is related to the level of internal anxiety or emotionality, and second, that the intensity of this anxiety could be ascertained by a paper and pencil test consisting of items describing what have been called overt or manifest symptoms of this state.

Since the scale has proved to be such a useful device in the selection of subjects for experimental purposes, a description of the construction of the test and the normative data that have been accumulated in connection with it may be of interest to other investigators in the field of human motivation.

DEVELOPMENT OF THE SCALE

The manifest anxiety scale was originally constructed by Taylor (6) for use in a study of eyelid conditioning. Approximately 200 items from the Minnesota Multiphasic Personality Inventory were submitted to five clinicians, along with a definition of manifest anxiety that followed Cameron's (2) description of chronic anxiety reactions. The judges were asked to designate the items indicative of manifest anxiety according to the definition. Sixty-five items on which there was 80 per cent agreement or better were selected for the anxiety scale. These 65 statements, supplemented by 135 additional "buffer" items uniformly classified by the judges as non-indicative of anxiety, were administered in group form to 352 students in a course in introductory psychology. The measures ranged from a low anxiety score of one to a high score of 36, with a median of approximately 14. The form of the distribution was slightly skewed in the direction of high anxiety.

Subsequently, the scale went through several modifications.[1] At present it consists of 50 of the original 65 items that showed a high correlation with the total anxiety scores in the original group tested. Furthermore, the buffer items have been changed so that the total test, which has been lengthened from 200 to 225 items, includes most of the items from the *L*, *K*, and *F* scales of the MMPI and 41 items that represent a rigidity scale developed by Wesley (10). The 50 anxiety items are reproduced in Table 1, along with the responses to these items considered as "anxious" and the ordinal numbers of the statements as they appear in the present form of the test.

Normative Data

Under the innocuous title of *Biographical Inventory*, the test in its present form has been administered to a total of 1971 students in introductory psychology at the State University of Iowa during five successive semesters from September, 1948 to June, 1951. The distribution for this sample is presented in Fig. 1. As can be seen by inspection, the distribution shows a slight positive skew, as did the original scale. The fiftieth percentile falls at about 13, the eightieth at about 21, and the twentieth at about 7. The mean of the distribution is 14.56.

[1] HEDLUND, J. L., FARBER, I. E., & BECHTOLDT, H. P. Normative characteristics of the Manifest Anxiety Scale. Unpublished paper. The statistical analysis, along with most of the data collected with the scale, was carried out under the direction of H. P. Bechtoldt at the State University of Iowa.

Sex Differences

A comparison of the scores of males and females in this total sample revealed that the mean score of the women was somewhat higher. The difference between the two means however was not statistically significant. For this reason, both sexes have been included in a single distribution.

Consistency of Scores

In order to determine the stability of the anxiety scores over time, groups of individuals have been retested on the scale after various intervals. In one instance, the results of retesting 59 students in introductory psychology after a lapse of three weeks yielded a Pearson product-moment coefficient of .89.

TABLE 1

ITEMS INCLUDED ON THE MANIFEST ANXIETY SCALE AND RESPONSES SCORED AS "ANXIOUS" ITEMS ARE NUMBERED AS THEY APPEAR IN THE COMPLETE BIOGRAPHICAL INVENTORY

4. I do not tire quickly. (False)
5. I am troubled by attacks of nausea.* (True)
7. I believe I am no more nervous than most others.* (False)
11. I have very few headaches. (False)
13. I work under a great deal of tension.* (True)
14. I cannot keep my mind on one thing. (True)
16. I worry over money and business. (True)
18. I frequently notice my hand shakes when I try to do something. (True)
24. I blush no more often than others.* (False)
25. I have diarrhea once a month or more.* (True)
26. I worry quite a bit over possible misfortunes.* (True)
27. I practically never blush. (False)
33. I am often afraid that I am going to blush. (True)
35. I have nightmares every few nights. (True)
36. My hands and feet are usually warm enough. (False)
37. I sweat very easily even on cool days. (True)
38. Sometimes when embarrassed, I break out in a sweat which annoys me greatly.* (True)
41. I hardly ever notice my heart pounding and I am seldom short of breath.* (False)
43. I feel hungry almost all the time. (True)
44. I am very seldom troubled by constipation.* (False)
48. I have a great deal of stomach trouble. (True)
51. I have had periods in which I lost sleep over worry.* (True)
54. My sleep is fitful and disturbed.* (True)
56. I dream frequently about things that are best kept to myself.* (True)
66. I am easily embarrassed. (True)
67. I am more sensitive than most other people.* (True)
77. I frequently find myself worrying about something.* (True)
82. I wish I could be as happy as others seem to be.* (True)
83. I am usually calm and not easily upset. (False)
86. I cry easily. (True)
87. I feel anxiety about something or someone almost all the time.* (True)
94. I am happy most of the time. (False)
99. It makes me nervous to have to wait. (True)
100. I have periods of such great restlessness that I cannot sit long in a chair.* (True)
103. Sometimes I become so excited that I find it hard to get to sleep. (True)
107. I have sometimes felt that difficulties were piling up so high that I could not overcome them.* (True)
112. I must admit that I have at times been worried beyond reason over something that really did not matter.* (True)
117. I have very few fears compared to my friends.* (False)
123. I have been afraid of things or people that I know could not hurt me. (True)
136. I certainly feel useless at times. (True)
138. I find it hard to keep my mind on a task or job. (True)
145. I am unusually self-conscious.* (True)
152. I am inclined to take things hard.* (True)
153. I am a high-strung person.* (True)
163. Life is a strain for me much of the time.* (True)
164. At times I think I am no good at all. (True)
168. I am certainly lacking in self-confidence.* (True)
183. I sometimes feel that I am about to go to pieces.* (True)
187. I shrink from facing a crisis or difficulty.* (True)
190. I am entirely self-confident.* (False)

* Statements rewritten for subsequent revision.

Different Populations

Scores on the scale are also available for samples drawn from somewhat different populations. Distributions for 683 airmen tested at the beginning of basic training at Lackland Air Force Base and for 201 Northwestern University night-school students of introductory psychology show essentially the same form as the group reported above, while the quartiles are in close agreement.

In a second test-retest study,[2] the scale was given to 163 students in an advanced undergraduate psychology course who had previously taken the test as introductory students. For 113 of these cases 5 months had elapsed since the first testing, while an interval of 9–17 months had intervened for the remaining 50. The test-retest coefficient was found to be .82 over 5 months and .81 for the longer

[2] See footnote 1.

period. Furthermore, no systematic change, upwards or downwards, was found in these distributions, i.e., the means of each of the three sets of scores remained essentially the same after retesting. Thus, for all groups tested, both the relative position of the individual in the group and his absolute score tended to remain constant over relatively long periods of time.

tered to 282 freshmen males, and approximately 18 weeks later the group MMPI was given to the same students. The correlation between the two sets of measures, obtained by determining the scores on the 50 anxiety items on each test, was .68. This, it will be noted, is a slightly lower figure than that obtained by test-retest on the Inventory after a comparable length of time. In addition,

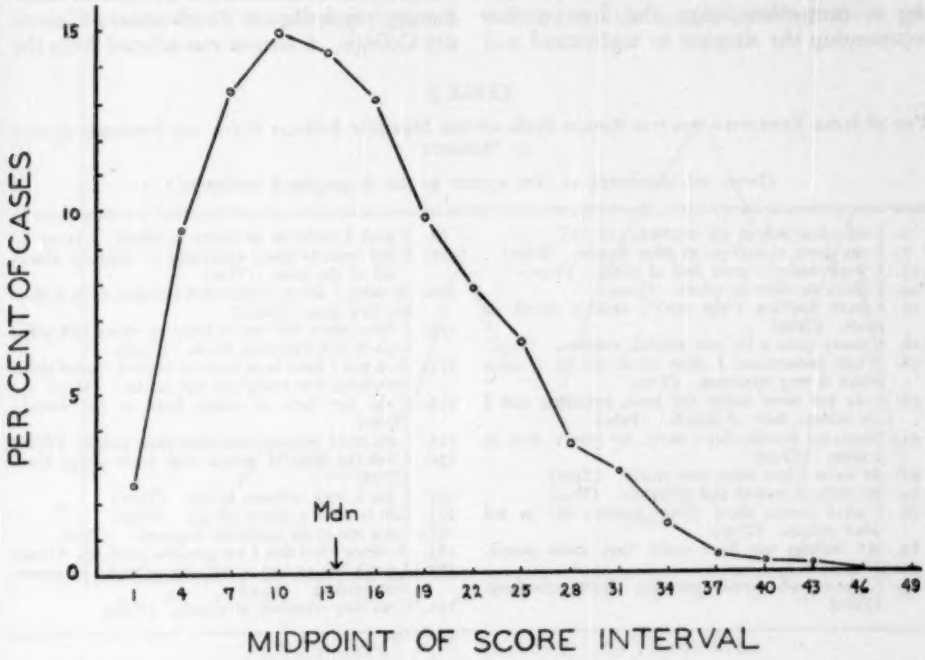

Fig. 1. Frequency Polygon Showing Per Cent of the 1971 University Students Receiving the Indicated Scores on the Manifest Anxiety Scale

Relationship of the Biographical Inventory to the MMPI

Since it might be desired to obtain anxiety scores for individuals who have been given the complete MMPI rather than the Biographical Inventory, it is necessary to consider the effects of the different sets of filler items on the 50 anxiety statements. There is some evidence [3] to suggest that the distribution of anxiety scores given in the form of the MMPI will differ significantly from that obtained from the Biographical Inventory. The Biographical Inventory was administered to 282 freshmen males, and approximately 18 weeks later the group MMPI was given to the same students. The correlation between the two sets of measures, obtained by determining the scores on the 50 anxiety items on each test, was .68. This, it will be noted, is a slightly lower figure than that obtained by test-retest on the Inventory after a comparable length of time. In addition, the forms of the distributions were statistically different, as indicated by a chi-square test of homogeneity. Since the initial scores of this group, obtained from the Biographical Inventory, were similar to those found with other groups, the discrepancy of the results between the Inventory and the MMPI suggests that the radical change in filler items may exert a definite influence on the anxiety scores. Before anxiety scores obtained from the MMPI can be evaluated it would appear to be necessary to have more normative data concerning the scale scores obtained from this form.

[3] See footnote 1.

REVISION OF THE SCALE

A further revision of the scale is now being carried out by the writer. This variation represents an attempt to simplify the vocabulary and sentence structure of some of the anxiety items that appear to be difficult to comprehend, especially for a noncollege population. Toward this end, the 50 anxiety items were first submitted to 15 judges who were instructed to sort them into four piles according to comprehensibility, the first position representing the simplest to understand and the fourth the most difficult. It was found that 28 of the items had a mean scale value of 2.00 or more. These 28 items were selected for revision and rewritten in at least two alternate forms.[4] Each set of alternatives was then ranked by a different set of 18 judges, first for ease of understanding and then for faithfulness of meaning to the original statement. For most of the items, the alternative judged to be simplest was also chosen as being closest in meaning to the original item and was therefore selected for the new scale. For those items in which a discrepancy occurred, faithfulness of meaning was chosen over simplicity. However, in every case, the new statement selected for inclusion on the scale was judged simpler than the original. These 28 rewritten items are shown in Table 2.

[4] In rewriting the items, the Thorndike word count (9) was consulted. These counts primarily determined substitution of words within an item whenever this was done.

Relationship between the Old and New Versions of the Scale

To demonstrate the relationship between the old and new versions of the test, both forms were administered to students in introductory psychology at Northwestern University College. A sample was selected from the college population for this purpose since it was thought that this group would show the least confusion in interpreting the original versions of the difficult items and, therefore, better demonstrate the comparability of the two forms than less verbally sophisticated individuals. Scores obtained from 59 students showed a Pearson product-moment correlation of .85 between the old and new versions, the latter being administered three weeks after the initial testing. This figure is quite comparable to the test-retest coefficient found for the previous form of the scale after a similar time interval. Considering only the 28 rewritten items, the correlation becomes .80.

While the correlation coefficient shows the high degree of relationship between the old

TABLE 2

THE 28 ITEMS REWRITTEN FOR THE REVISED FORM OF THE MANIFEST ANXIETY SCALE AND RESPONSES SCORED AS "ANXIOUS"

(Items are numbered as they appear in the Biographical Inventory.)

5. I am often sick to my stomach. (True)	82. I wish I could be as happy as others. (True)
7. I am about as nervous as other people. (False)	87. I feel anxious about something or someone almost all of the time. (True)
13. I work under a great deal of strain. (True)	100. At times I am so restless that I cannot sit in a chair for very long. (True)
24. I blush as often as others. (False)	107. I have often felt that I faced so many difficulties I could not overcome them. (True)
25. I have diarrhea ("the runs") once a month or more. (True)	112. At times I have been worried beyond reason about something that really did not matter. (True)
26. I worry quite a bit over possible troubles. (True)	117. I do not have as many fears as my friends. (False)
38. When embarrassed I often break out in a sweat which is very annoying. (True)	145. I am more self-conscious than most people. (True)
41. I do not often notice my heart pounding and I am seldom short of breath. (False)	152. I am the kind of person who takes things hard. (True)
44. Often my bowels don't move for several days at a time. (True)	153. I am a very nervous person. (True)
51. At times I lose sleep over worry. (True)	163. Life is often a strain for me. (True)
54. My sleep is restless and disturbed. (True)	168. I am not at all confident of myself. (True)
56. I often dream about things I don't like to tell other people. (True)	183. At times I feel that I am going to crack up. (True)
67. My feelings are hurt easier than most people. (True)	187. I don't like to face a difficulty or make an important decision. (True)
77. I often find myself worrying about something. (True)	190. I am very confident of myself. (False)

and revised forms, the question still remains as to whether rewriting the 28 items has reduced the difficulty level of these statements so as to minimize confusion and misinterpretation. In an attempt to determine this, the scores of the 59 students given both versions were analyzed into two components: that for the 28 difficult items and that for the 22 items left intact. For each form, scores on the 28 items were correlated with the remaining 22. It was reasoned that if the original forms of the 28 items were confusing, then the rewritten items, if attempts to simplify were successful, would show a higher correlation with the 22 items left intact than would the original statements. The actual correlations obtained in this manner were .81 for the old version and .83 for the new. Although the difference between the coefficients was in the desired direction, a *t* test indicated that it was statistically insignificant. However, a significant difference in correlations might be obtained with subjects of lesser educational attainment since misinterpretation of the 28 original items would be more likely to occur with such a group.

Normative Characteristic of the New Scale

To determine further characteristics of the distribution of scores on the new version, 229 students in introductory psychology were given only the revised form of the scale (1). It was found that the shape of the distribution and the values of the quartiles did not differ significantly from those obtained with the previous form.

Retest scores are also available for 179 individuals from the sample described above. A product-moment correlation of .88 was found after an intertest interval of four weeks. However, while the position of the individuals in the group tended to remain the same, a downward shift in the absolute scores of the entire distribution was noted from test to retest. The difference between means (14.94 vs. 12.92) was significant at the .01 level of confidence, as indicated by a *t* test.

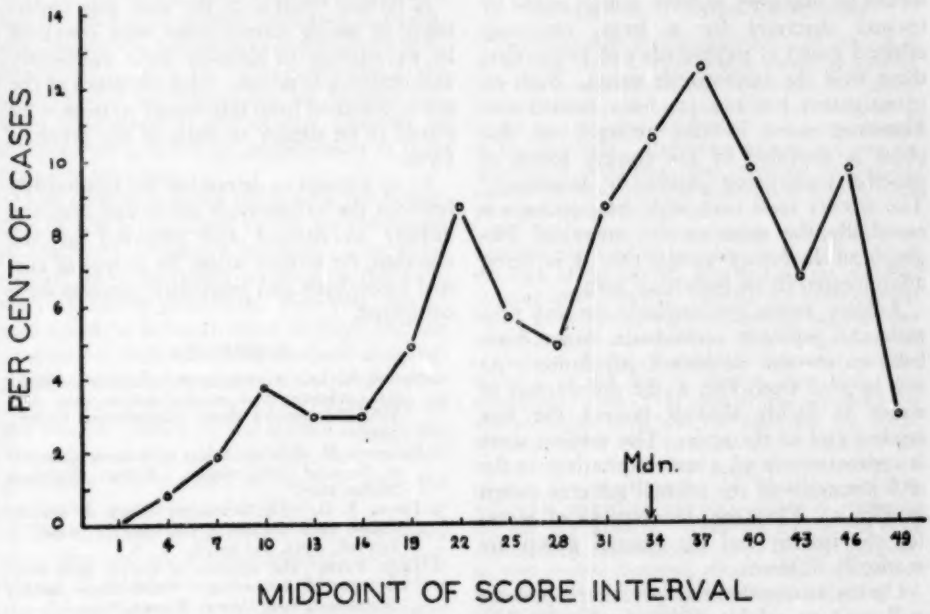

Fig. 2. Graph of the Frequency Distribution of Manifest Anxiety Scores Received by 103 Psychiatric Patients

Relationship of the Anxiety Scale to Other Measures

The anxiety scale was developed for, and has been used exclusively as, a device for selecting experimental subjects, without re-

gard to the relationship of the scores to more common clinical definitions (e.g., clinical observation). While defining degree of anxiety in terms of the anxiety-scale scores is a perfectly legitimate operational procedure, determining the relationship between this definition and clinical judgments might extend the applicability of both the scale and the experimental results found in the studies utilizing the scale.

In order to determine the relationship between the scale and clinical judgments, it would be necessary to have ratings made by trained observers for a large, randomly selected group of individuals and to correlate these with the anxiety-scale scores. Such an investigation has not yet been carried out. However, some indirect evidence on this point is provided by the anxiety scores of patients undergoing psychiatric treatment.[5] The anxiety scale used with these patients is essentially the same as the unrevised Biographical Inventory except that it is being administered in an individual form.

Anxiety scores are available for 103 neurotic and psychotic individuals, drawn from both an in- and outpatient population. As can be seen from Fig. 2, the distribution of scores is highly skewed toward the low anxiety end of the scale. The median score is approximately 34, a score equivalent to the 98.8 percentile of the normal subjects shown in Fig. 1. Thus the distributions of scores for the patient and the normal group are markedly different.

On the assumption that psychiatric patients will tend to exhibit more manifest anxiety symptoms (as determined by direct observation) than do normal individuals, this difference between the two groups appears to indicate that there is some relationship between the anxiety-scale scores and clinical observation of manifest anxiety.

SUMMARY

A manifest anxiety scale, consisting of items drawn from the Minnesota Multiphasic Personality Inventory judged by clinicians to be indicative of manifest anxiety, was developed as a device for selecting subjects for experiments in human motivation.

After statistical analysis the original 65-item scale was reduced to the 50 most discriminating statements. These items, supplemented by 225 statements nonindicative of anxiety, are given under the title of the Biographical Inventory. Normative data and test-retest correlations found with scale scores taken from the Biographical Inventory are presented.

A further revision of the scale was undertaken in which certain items were rewritten in an attempt to simplify their vocabulary and sentence structure. Characteristics of the scores obtained from this revised version were found to be similar to those of the previous form.

In an attempt to determine the relationship between the anxiety-scale scores and manifest anxiety as defined and observed by the clinician, the anxiety scores for groups of normal individuals and psychiatric patients were compared.

[5] These data are obtained from a study currently being conducted by the writer and K. W. Spence investigating the role of anxiety in neurotic and psychotic disorders by means of an eyelid conditioning technique.

REFERENCES

1. AHANA, ELLEN. A study on the reliability and internal consistency of a manifest anxiety scale. Unpublished master's thesis, Northwestern Univer., 1952.
2. CAMERON, N. *The psychology of behavior disorders: a bio-social interpretation.* Boston: Houghton Mifflin, 1947.
3. LUCAS, J. D. The interactive effects of anxiety, failure, and intraserial duplication. *Amer. J. Psychol.*, 1952, 65, 59–66.
4. PECK, RUTH. The influence of anxiety upon effectiveness of counseling. Unpublished doctor's dissertation, State Univer. of Iowa, 1950.
5. SPENCE, K. W., & TAYLOR, JANET. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, 42, 183–188.
6. TAYLOR, JANET A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, 41, 81–92.
7. TAYLOR, JANET A., & SPENCE, K. W. The relationship of anxiety to performance in serial learning. *J. exp. Psychol.*, 1952, 44, 61–64.
8. THORNDIKE, E. L., & LORGE, I. *Teacher's word book of 20,000 words.* New York: Teachers College, Columbia Univer., 1941.
9. WENAR, C. Reaction time as a function of manifest anxiety and stimulus intensity. Unpublished doctor's dissertation, State Univer. of Iowa, 1950.
10. WESLEY, ELIZABETH L. Perseverative behavior in a concept formation task. Unpublished doctor's dissertation, State Univer. of Iowa, 1950.

Received June 26, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

TRAUMATIC AVOIDANCE LEARNING: THE OUTCOMES OF SEVERAL EXTINCTION PROCEDURES WITH DOGS

RICHARD L. SOLOMON, LEON J. KAMIN, AND LYMAN C. WYNNE[1]
Harvard University

THE persistence of behavior with no obvious reinforcement poses special problems for theories of learning. But numerous experimental studies confirm the clinical observation that traumatically acquired habits maintain a marked resistance to extinction despite the lack of renewed primary reinforcement (3, 4, 6, 8). The generality of the problem has been indicated by Mowrer (9), who formulates as the central problem of neurosis the paradoxical perpetuation of nonadaptive behavior.

The present research deals with the extinction of a habit acquired under the conditions of traumatic avoidance learning. The habit proved to be markedly resistant to the ordinary extinction procedure. However, extinction could be brought about through the use of special techniques. The obtained data suggested theoretical explanations of high resistance to extinction as well as explanations of the relative efficacy of the special techniques. We presume that the findings have a wide generality for theories of therapy and the nature of psychological trauma.

The findings to be reported in this paper were obtained under the following conditions. Using a modification of the Mowrer-Miller shuttlebox, learned avoidance responses were established in normal mongrel dogs. Under the impetus of very intense (just subtetanizing) electric shock, the dogs were trained to jump over a barrier from one compartment to the other in order to avoid the shock. The details of the training procedure are given in another paper (13). Ten trials comprised each day's session. After the animals had met an acquisition criterion of 10 successive responses on a given day, different dogs were subjected to different types of extinction procedures. It is the results of a series of such extinction experiments which will be the subject of this paper.

[1] This research was supported by grants from the Medical Sciences Division of the Rockefeller Foundation. It was facilitated by the Laboratory of Social Relations, Harvard University. Dr. Wynne was a U.S. Public Health Service Postdoctoral Research Fellow in Mental Health when this work was initiated. The authors wish to thank E. S. Brush, F. R. Brush, N. Kogan, B. N. Cohn, and E. Smulekoff for research assistance and many useful suggestions.

EXPERIMENT 1. PILOT EXPERIMENT

Two dogs were trained with a CS-US interval of 10 seconds, with a three-minute intertrial interval. The CS consisted of two segments: a buzzer sounded for one second, together with the raising of the gate which separated the two compartments of the shuttlebox. After these animals had met the criterion of acquisition, they were not shocked regardless of how long they remained in a given compartment without jumping.[2]

We had expected these dogs to extinguish spontaneously since the barrier was set at the height of the back of the animal, thus making the jumping response quite effortful. Instead, the experimenters found themselves running the animals day after day with no signs of extinction. Indeed, the latencies of response to the CS gradually *decreased* and the behavior of the animals became more and more stereotyped. One of the animals was continued for 190 extinction trials and the other for 490 trials. The dog which continued for 490 extinction trials had received only 11 shocks during the acquisition phase. The experimenters felt that this ordinary extinction procedure was *not efficient* in removing the jumping response. While the dogs *might* have extinguished spontaneously after several hundred, or even several thousand, trials more, the behavior and latencies of the animals gave no suggestion of this conceivable eventuality.

An attempt was then made to discourage the 490-trial dog from jumping by electrifying the opposite compartment on each trial, so that the dog jumped *into* shock. The gate was immediately lowered after each jump to prevent retracing, and the shock, at just subtetanizing level, was continued for three seconds and then terminated. The dog became more upset, and at subsequent presentations of the CS jumped more vigorously. His latencies were maintained at their already extremely fast level of 1.0–1.2 seconds. After 100 additional trials under this shock-extinction procedure, the dog was still jumping regularly into shock and gave no signs of extinguishing. As he jumped on each trial, he gave a sharp anticipatory yip which turned into a yelp when he landed on the electrified grid in the

[2] The reader is referred to another paper of this series (13) for precise details of the training procedure and the consequent behavior of the dogs during the *acquisition phase* of avoidance learning.

opposite compartment of the shuttlebox. We then increased the duration of the shock to 10 seconds and ran the dog for 50 more trials. Latencies and behavior did not change. Evidently punishment for jumping was ineffective in extinguishing the jumping response.

We tried a third extinction procedure with the same dog, one designed to prevent the dog from jumping in the presence of the CS. We placed a plate of glass in the opposite compartment, flush against the barrier and gate so that it blocked the passage between the two compartments. The plate of glass could not be seen by the dog until the gate was raised and the buzzer was sounded. The glass barrier was inserted on trials 4–7 of each extinction day. On the other six trials, the dog was shocked for jumping as before. This procedure was intended to have a "reality-testing" function for the dog. Ostensibly, the procedure let the dog "know" that the presence of the CS no longer signified shock. On the first glass-barrier trial, he jumped forward immediately at the CS and smashed his head against the glass. He drew back and was fairly quiet, and the gate was lowered after two minutes. On the subsequent trials, on which he did not strike the glass, he barked furiously, panted very rapidly, quivered, and drooled while the CS was present, but quieted down after the gate was lowered, and remained quiet during the minute before the next trial was started. On the 8th–10th trials, when the glass barrier was no longer present, he did not attempt to jump, and the gate was lowered after two minutes. This two-minute period constituted the arbitrary criterion for no response. At long last, after 647 extinction trials, the dog failed to jump in the presence of the CS alone. On the following day, the dog jumped into shock with short latencies on the first three trials, thus showing complete spontaneous recovery. He did not jump on trials 8–10. On five subsequent days, during which the glass barrier was never present, he jumped only on the second trial of the third day with a latency of 1.3 seconds. With this dog there was no gradual lengthening of latencies, but, rather, extinction was an all-or-none affair.

Two more animals were run with the same training procedures as with the first two, except that 20 trials a day, instead of 10, were carried out. The ordinary extinction procedure was used. One animal was discontinued after 310 extinction trials and the other after 280 extinction trials. During these trials, neither animal showed any signs of spontaneous extinction.

Some of the questions posed by the results of this pilot study suggested the succeeding experiments: (*a*) Did the shock-extinction procedure (punishment) fail because the animal had so much practice in jumping before the procedure was instituted? That is, *does the ordinary extinction procedure decrease the likelihood of effectiveness of the special procedures?* (*b*) Would the "reality-testing," glass-barrier procedure have been effective if it had been used directly after ordinary extinction? (*c*) Would the combined glass-barrier and shock-extinction procedure have been as effective if it had been preceded by the glass-barrier procedure rather than by shock extinction?

EXPERIMENT 2. THE EFFECTS OF ORDINARY EXTINCTION PROCEDURE

This experiment was designed to test the efficacy of the ordinary extinction procedure. The simplest hypothesis might claim that the removal of all primary reinforcement (shock) should lead to either the gradual or sudden extinction of the jumping response.

Thirteen dogs were given avoidance training in the shuttlebox situation, using the following procedures. The CS consisted of turning off a light over the compartment in which the dog happened to be and raising the gate. The light stayed off and the gate stayed up until the dog jumped into the lighted compartment, after which the gate was lowered. The light-out signal was a modification of the procedure used in the pilot experiment, and it replaced the buzzer. As in the pilot experiment, an intense shock was used. The CS-US interval was 10 seconds, and the barrier set for each dog was at the height of his back. The intertrial interval was three minutes, and 10 trials were given each day. The criterion for acquisition of the jumping response was 10 avoidances in 10 trials. A complete description of the training procedures is given in another paper (13).

After the criterion of learning was reached, no shock was administered in the presence of the CS, no matter how long a dog might delay in jumping. Such a procedure, where the shock reinforcement is no longer administered, we shall call *ordinary extinction*. Ordinary extinction was carried on for 20 days after each dog met the learning criterion. Thus each dog received 200 trials following the criterion trials, or at least 210 trials since the last shock had been received during the acquisition phase.

During the ordinary extinction procedure, not a single animal delayed jumping long enough to meet the (arbitrary) criterion of no-response, or infinite latency, which was set at *two minutes without jumping in the presence of the CS*. In other words, no extinction was obtained. In fact, there were only 11 responses with a latency greater than 10 seconds during the 2582 extinction responses.[3] Rather than showing any signs of

[3] One dog missed 10 trials because of an oversight, one dog missed eight trials because of a sore foot; this accounts for the fact that the total number of responses was less than 2600 (20 days×10 trials×13 animals).

extinction, there was a general tendency for the latencies to shorten. Figure 1 presents a graph of the mean reciprocal of response latency as a function of days after meeting the criterion of learning. Each point represents the mean for the 13 dogs for 10 trials for a given day of extinction. The arrow at the origin of the abscissa designates the mean for all dogs for the 10 criterion trials. It can be seen that the dogs were jumping with a mean latency of approximately 2.7 seconds during the criterion trials. The mean latencies gradually decreased, and after 200 ordinary extinction trials, the dogs were jumping with a mean latency of approximately 1.6 seconds.

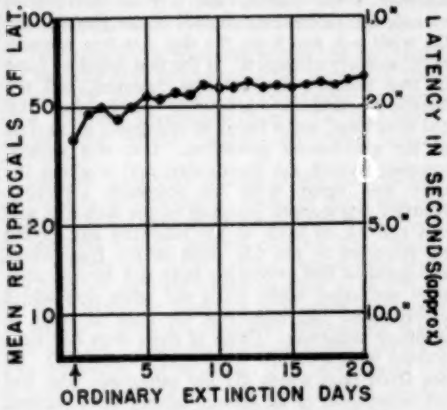

FIG. 1. MEAN RECIPROCALS OF LATENCY AS A FUNCTION OF NUMBER OF ORDINARY EXTINCTION DAYS

Each point is the mean reciprocal of response latency for thirteen dogs for the ten trials of a given day. Note that the ordinate for mean latency in seconds is approximate because the antireciprocal of the mean reciprocal of latency is not necessarily the same as the mean latency in seconds.

An inspection of the curve suggests that the latency asymptote had not yet been reached after 200 trials. This is not surprising, since it will be remembered that one of the dogs in the pilot experiment was jumping regularly at 1.0–1.2 seconds at the end of 490 extinction trials.

During the extinction trials, and accompanying the decrease in latency of the jumping response, the experimenters were impressed with certain behavioral changes. First, the behavior of the dogs, both in style of jumping and in intertrial behavior, became stereotyped. Each dog appeared to have developed his own particular "ritual." Such rituals involved actual response patterns as well as orientation in the apparatus. For example, a protocol description of one dog's behavior reads as follows: "At CS responded immediately and made clean jump with no pause. Then sat immobile in rear-door corner, with his back six inches from the rear, along wall opposite door, facing barrier gate, but nose pressed against wall. No panting or other reactions." For the next trial, the behavior was precisely the same, with the position again along the left side of the compartment with reference to the barrier, nose pressed against the wall.

Second, the frequency and intensity of overt emotional reactions, both to the CS and in the intertrial interval, decreased markedly. The types of emotional signs which usually disappeared during the course of ordinary extinction were defecation, urination, yelping and shrieking, trembling, attacking the apparatus, scrambling, jumping on the walls of the apparatus, and pupillary dilation. Whining, barking, and drooling tended to decrease in magnitude but often persisted throughout the 200 trials. Panting tended to persist relatively undiminished if it occurred during the early trials. Some dogs showed no overt emotional signs during the latter part of ordinary extinction. All dogs, early in the extinction procedure, showed some resistance—often very strenuous—to being placed into the apparatus. But most dogs, after 10 or 12 days of extinction, no longer resisted being placed into the apparatus. Many voluntarily hopped inside, displaying no visible emotional response.

Not only were there no signs of extinction in our dogs from day to day, but there was very little change in mean latency from trial to trial within experimental days. In Figure 2 are plotted the mean latencies for all dogs as a function of trials-within-days. On the first trial of each day, the mean reciprocal of latency for the 20 days was 56.1, while the mean reciprocal for the tenth trial was 55.1. This change is insignificant, representing but a small fraction of a second. There appears to be a slight warm-up effect on the first three trials, but this trend is also

insignificant. Spontaneous recovery from the last trial of one day to the first trial of the next is minute, if present at all; it is indicated by the dotted line. The slight downward trend of the reciprocals between the third and tenth trials might indicate that longer experimental sessions would be more conducive to producing longer latencies, but this is doubtful.

The results of this experiment indicate that the ordinary extinction procedure is quite ineffective for eliminating the jumping response. If anything, it seems only to strengthen it. It can of course be argued that

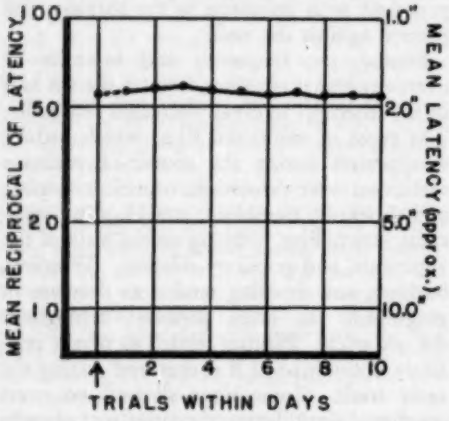

FIG. 2. MEAN RECIPROCALS OF LATENCY AS A FUNCTION OF TRIALS WITHIN DAYS

Each point is the mean reciprocal of response latency for thirteen dogs for twenty days of ordinary extinction procedure. Note that the ordinate for latency in seconds is approximate.

if the procedure had been continued for an indefinitely long time extinction *might* have occurred. But we must stress that within the time spans covered by this and the pilot experiment there were *no* signs of extinction of jumping.

EXPERIMENT 3. THE EFFECTS OF THE GLASS-BARRIER, REALITY-TESTING PROCEDURE

This experiment was designed to test the efficacy of an extinction procedure which physically prevented the animal from jumping in the presence of the CS. From one point of view, this technique corresponds to "reality testing." The animal was forcibly exposed to a sequence of events in which the CS, followed by *not*-jumping, is *no longer* followed by shock. Will the dog, as a result of this new "knowledge," stop jumping?

Nine dogs were trained in the same manner as those of Experiment 2, with the exception that four had learned with a CS-US interval of 20 seconds instead of 10 seconds. Some of these dogs had been used in Experiment 2. Five of the dogs were carried through the 200 trials of ordinary extinction, just as in Experiment 2, while the other four were given only 10 ordinary extinction trials beyond the 10 criterion trials. Thus, we had a group of animals which had jumped approximately 10 times as often as the other group. Then the glass-barrier procedure was introduced as follows: On trials 4–7 of a given day, a plate of glass was placed flush against the barrier and gate on the side opposite the compartment containing the dog. The glass had three narrow vertical strips of adhesive tape placed upon it. After presentation of the usual CS, these adhesive strips enabled the dog to discriminate visually the presence or absence of the glass-barrier. On trials 1–3, and 8–10, the dog was free to jump as in ordinary extinction. If the dog failed to jump in two minutes, the CS was removed, and the latency was defined as infinite or "no-response." Of the nine dogs, seven failed to extinguish in 10 days of the glass-barrier procedure. One dog stopped jumping during the fourth day, and was run the next day again with no responses occurring. Another dog stopped jumping on the fifth day, and was run for 10 trials on the following day, giving no responses to the CS. One of the dogs which extinguished had previously been run for 200 ordinary extinction trials, while the other dog which extinguished had been run for only 20 trials of ordinary extinction. (Both of these dogs had been trained with a CS-US interval of 10 seconds.) Of the seven dogs which *did not* extinguish, four had no occurrences of an infinite latency (two minutes without jumping). Only on trials 8–10 were there any infinite latencies for the other three animals. In most cases where infinite latencies, or very long ones, occurred on the eighth trial, the animals behaved in very much the same way as they did when the glass barrier was present. The animals looked as though they "thought" the glass barrier was still there. In two instances where trial 10 produced an infinite latency, on the following day the first trial latencies were 1.3 seconds and 1.5 seconds, respectively. These were dramatic cases of spontaneous recovery. Of the seven animals which did not extinguish, only two had any latencies over 5 seconds on trials 1–3. Of the nine animals in the experiment, eight showed an all-or-none pattern whenever the latencies changed. That is, the animals either jumped quickly (from 1 to 4 seconds) at the presentation of the CS, or they did not jump for long time periods, in some cases not at all. Such long latencies were then typically followed by very short ones on the next trial. Such all-or-none changes were also noted in our study of the characteristics of acquisition (13). There was no gradual lengthening of latencies during the 10 glass-barrier days. Even the two dogs which extinguished showed a sudden transition from short

latencies to infinite ones during the course of their extinction. Our protocols indicated that most dogs, when first confronted with the glass barrier, exhibited a wide variety of intense emotional responses. These usually disappeared after several repetitions of the glass-barrier procedure.

The results of this experiment indicate that the glass-barrier procedure *does* produce extinction in *some* dogs. However, *within the ten-day span covered by the experiment,* the procedure failed to extinguish the jumping response in most of the dogs. It is to be noted that there was no difference between dogs trained for 10 and for 200 ordinary extinction trials, one of each group extinguishing.

EXPERIMENT 4. THE EFFECTS OF THE SHOCK-EXTINCTION, PUNISHMENT PROCEDURE

This experiment was designed to test the efficacy of an extinction procedure which punished (with shock) the performance of the jumping response. The simplest expectation might be that shock-punishment would "stamp out" the jumping response.

Thirteen dogs were trained in the same manner as those in Experiment 2. Some of them had participated in Experiment 2. Seven were trained with a 10-second CS-US interval and six were trained with a 20-second CS-US interval. Seven were given 200 trials of ordinary extinction, just as in Experiment 2, while the other six dogs were given only 10 ordinary extinction trials after meeting the learning criterion. Then the shock-extinction procedure was introduced for all dogs. This was carried out as follows: The dog was shocked for three seconds in the compartment *into which* he jumped. No shock was administered if the animal did not jump in the presence of the CS. The immediacy of the shock-for-jumping was guaranteed by having the grid onto which the animal jumped electrified before the presentation of the CS. The shock level was the same as that used in training. The gate was immediately lowered after a jump in order to prevent retracing. A two-minute latency was again arbitrarily defined as infinite, or "no-response," and the CS was removed at the end of this two-minute period.

Of the 13 dogs in this experiment, 10 failed to extinguish in 100 shock-extinction trials. Three dogs extinguished; one of these animals had received 200 trials of ordinary extinction and two had received 10 trials of ordinary extinction prior to the introduction of the shock extinction procedure. All these three gave infinite latencies in the first 11 trials of punishment for jumping. They all met a criterion of no responses in 10 trials at the end of the second day of shock extinction. There was *no spontaneous recovery* once an infinite latency had occurred: if one of these three animals failed to jump he did not jump again. The transition from jumping to not-jumping was abrupt in two of the three animals. There was a considerable lengthening of latencies on the trials which preceded the onset of infinite latencies for one of these dogs.

The ten animals which did *not* extinguish showed an *entirely different course of behavior.* Nine of the 10 exhibited *no* infinite latencies. In fact, there was a tendency for the latencies to shorten on the first shock-extinction trials. These dogs jumped *faster* and *more vigorously* into the shock than they had jumped previously under the ordinary extinction procedure. (See Table 1 for a description and analysis of these data.) They often slammed into the far end of the compartment into which they were jumping. In addition, they all developed anticipatory reactions prior to jumping, which indicated that they "knew" they were to be shocked. The most common reaction was to yelp at the CS, jump vigorously, and then yelp at the shock, barking rapidly when the shock was terminated. This behavior continued for 100 trials of the shock-extinction procedure. Some long latencies did occur, but they were scattered and were usually followed by short latencies. However, one dog had six infinite latencies on the second day of shock extinction. Yet this dog failed to extinguish, and his latencies on the tenth day were short!

Several types of trend were evident in the latencies of the 10 dogs which did *not* extinguish. It is impossible, therefore, to generalize for all animals. Some dogs showed progressively shortening latencies, and some showed a slight and gradual lengthening of latencies. None showed radically lengthened latencies. One dog showed lengthening latencies for the first three days of shock extinction and thereafter showed gradually shortening latencies. Despite these different trends, trial-to-trial and day-to-day variability in latencies was very small during the 100 trials of shock extinction *for those animals which did not extinguish.* All animals were extremely upset by the procedure, exhibiting symptoms of terror.

The results of this experiment indicate that

the shock-extinction procedure *does* produce abrupt extinction in *some* dogs. However, within the ten-day span covered by the experiment, the procedure failed to extinguish the jumping response in most of the dogs. Punishment seemed to *increase* the strength of the jumping response in most dogs, as indicated by shorter latency and greater vigor of response. Again, there is no significant difference between animals trained for 10 and for 200 ordinary extinction trials.

if the dog jumped in the presence of the CS he jumped into shock. The shock was on for three seconds. The gate was lowered to prevent retracing. (If the dog did not jump in two minutes, the CS was withdrawn and the trial was scored as infinite in latency, or "no-response.") Then, on trials 4 through 7, the glass barrier was present, just as in Experiment 3. The CS was removed after two minutes' exposure to the glass barrier. Of course, the animal could not jump to the opposite compartment during trials 4–7. Then, on trials 8–10 the same procedure used on trials 1–3 was introduced again, and the animals were shocked if they jumped to the CS. Thus the shock-extinction procedure was used on the first and last three trials of each day, and the glass-barrier procedure was used on the middle four trials. It was decided arbitrarily to terminate this experiment at the end of 10 days if no extinction occurred.

TABLE 1

Mean Reciprocals of Response Latency for the Five Trials Preceding and Following the Onset of the Shock-Extinction Procedure

	Shock Extinction Follows 200 Ordinary Extinction Trials	Shock Extinction Follows Only 20 Ordinary Extinction Trials
Before shock extinction	50.7	50.5
After onset of shock extinction	72.9	63.1

Analysis of Variance

Source	*F*	*p*
A. Trials before and after onset of shock extinction	15.02	>.01
B. Number of trials preceding onset of shock-extinction procedure	1.21	—
A×B Interaction	2.50	—

Experiment 5. The Effects of a Combination of the Glass-Barrier Procedure and the Shock Extinction Procedure

This experiment was designed to test the efficacy of a procedure which combined in close juxtaposition the two special extinction techniques tested in the preceding experiments. Here punishment and reality-testing both occurred within the same day's experimental session.

Sixteen dogs were trained in the acquisition of avoidance in the same manner as those in Experiment 2. Of these, seven were trained with a 10-second CS-US interval and nine were trained with a 20-second CS-US interval. Ten had 200 trials of ordinary extinction, just as in Experiment 2, while the other six dogs were given only 10 trials of ordinary extinction after meeting the learning criterion. All of the dogs in this experiment had failed to extinguish in 100 trials of *either* the glass-barrier procedure *or* the shock-extinction procedure. Ten of the dogs had been subjected to the shock-extinction procedure and six had experienced the glass-barrier procedure. Then the combination of glass-barrier and shock-extinction procedure was introduced in the following manner: On the first three trials, the glass barrier was not present, but

Fourteen out of 16 dogs extinguished before 7 days of the combined procedure. The extinction criterion, as in previous experiments, was 10 infinite latencies on a given day. When a dog failed to respond on the first three trials of a day, the glass barrier was not used, and 10 trials were given in order to test for the occurrence of extinction. The remaining two animals failed to extinguish in 10 days, or 100 trials. One of them had had 200 ordinary extinction trials, and one had had only 10 ordinary extinction trials. Both had had 100 shock extinction trials prior to the combined procedures. One of the two dogs which did not extinguish in 10 days of the combined procedure showed no signs of extinction after 10 days. The other dog

showed lengthening latencies and one infinite latency on the last day of the combined procedure. *All* of the six dogs which had had 10 days of *glass-barrier trials prior to the combined procedure* met the extinction criterion in 30 trials or less, with a mean of 25.0 trials to meet the criterion of extinction. The eight dogs which had had *shock extinction,* and which *did* extinguish in the combined procedure, met the criterion for extinction in less than 70 trials of the combined procedure, with a mean of 43.8 trials.

The mean number of shocks required for extinction in the *combined* procedure was 8.2 shocks for the six dogs which had had the glass-barrier procedure and 14.3 shocks for the eight dogs which had had the shock-extinction procedure. This difference (14.3–8.2) is statistically significant ($p=.01$).

The results of this experiment indicate that the combined glass-barrier and shock-extinction procedure is highly effective in producing extinction of the jumping response. Again there is no difference between dogs trained for 10 and for 200 ordinary extinction trials. There are, however, several indications that this technique is more effective when preceded by the glass-barrier rather than by the shock-extinction technique. We have no data on how effective the present combined procedure would have been if introduced immediately after the ordinary extinction procedure, without the intervention of a special technique.

SUMMARY OF RESULTS

A summary of the findings of Experiments 2, 3, 4, and 5 is given in Table 2. From this table, several generalizations are possible: (*a*) The ordinary extinction procedure is ineffective. (*b*) Either the glass-barrier procedure or the shock-extinction procedure *can* produce extinction of jumping, but the proportions of animals extinguishing are small. There is little to choose between the two procedures. (*c*) The combined procedure is the most effective. It makes little difference whether it is preceded by 200 or by 10 ordinary extinction trials. There are, however, some indications that the combined procedure is more effective following the glass-barrier procedure than following the shock procedure. Not only is this trend detectable in the proportions of animals extinguishing with each sequence, but also the trend is clearly present in the data on (*a*) number of trials required to achieve the extinction criterion, and (*b*) number of shocks in the combined procedure required to produce extinction.

TABLE 2

PROPORTIONS OF ANIMALS EXTINGUISHING WITH THE VARIOUS EXTINCTION PROCEDURES

EXPERIMENT	SPECIAL PROCEDURE USED	AFTER 200 TRIALS ORDINARY EXTINCTION PROCEDURE	AFTER 10 TRIALS ORDINARY EXTINCTION PROCEDURE	TOTAL
2	None	0/13	0/10	0/23
3	Glass-barrier procedure only	1/5	1/4	2/9
4	Shock procedure only	1/7	2/6	3/13
5	Combined procedure, glass-barrier procedure first	4/4	2/2	6/6
5	Combined procedure, shock procedure first	5/6	3/4	8/10

THEORY

We had, in planning the extinction experiments, definite expectations about the relative effectiveness of the various extinction procedures. For the most part, these expectations were derived from the two-process theory outlined in another paper on the acquisition of traumatic avoidance learning (13).

We assumed that, in the course of acquisition, conditioned emotional reactions would be established through the process of classical conditioning, of CS-US contiguity in time. Further, we assumed that these emotional reactions would give rise to stimuli having drive properties. When an animal in our experiment was regularly avoiding the shock, he was in reality escaping from the CS, and each jump was, therefore, followed by a

reduction in the intensity of the emotional response elicited by the CS. Thus, not only was the jumping response reinforced by shock termination during the escape phase of acquisition, but also by anxiety reduction during the early avoidance trials. Theoretically, we would expect that the jumping response would be reinforced on each occasion that the CS aroused anxiety and the jumping response terminated or reduced the anxiety. However, once the animal is regularly avoiding the shock, the CS is no longer followed by the US, and the temporal contiguity necessary to reinforce the conditioned emotional responses is no longer present. This dissociation of the CS and US should, according to classical conditioning principles, lead to extinction of the emotional responses in the presence of the CS. If this should occur, there should then be no basis for the reinforcement of the instrumental jumping responses; there would be no anxiety to be reduced. The instrumental response should then gradually extinguish. These formulations would predict a long course of extinction for avoidance responses because the conditioned emotional reactions would have to be extinguished before the extinction of the instrumental response would start to occur. However, extinction of the instrumental response should occur.

With such a theory in mind, we expected that the learned jumping response would continue to decrease in latency several trials after the last shock had been received. As long as the conditioned emotional response had not been extinguished, each jump in the presence of the CS would be followed immediately by anxiety reduction. According to the law of effect, the bond between the CS and the jumping response should be consistently strengthened, with a correlated progressive shortening of response latency. This, in fact, did occur. In Experiment 2, the latencies shortened over a period of about 100 trials following the trial on which the last shock was received. ***But it is at this point that a real weakness in our interpretation appears.***

In a paper on the acquisition of traumatic avoidance responses we argued that the latency of the *conditioned emotional reaction* should be of the order of magnitude 1.5–2.5 seconds (13). But, after approximately 100 trials of ordinary extinction, the latency of the entire *jumping* response had in fact decreased to a mean value of about 1.7 seconds, and at the end of 200 trials of ordinary extinction the mean latency was about 1.6 seconds.

The instrumental jumping response *itself* usually has a latency of from 1.0 to 1.5 seconds at its asymptote. (Some individual animals produced latencies as short as 0.9 seconds, though very rarely.) Thus, if the animals' jumps were responses to their own emotional reactions, or if they were responding to drive arousal, the asymptote for the latency of the jumping response in the presence of the CS should have been approximately 2.4 to 3.4 seconds, i.e., the sum of the latencies for drive arousal and for jumping. But the animals were jumping faster than this within 10 to 20 trials after meeting the acquisition criterion (see Fig. 1). Obviously, the animals at this point *could not* have been responding to their own emotional reactions. The assumption that autonomic arousal serves as the stimulus for the jumping response is contradicted by the facts. The latency data indicate that, at this point, the jumping response has an extremely high habit strength and is activated by minimal drive. In any event, it cannot be argued that the jumping is energized by a ***full-blown anxiety reaction to the CS.***

It is clear that either something was fundamentally wrong with our interpretation of acquisition (13), or that additional principles have been overlooked. For, once the animals are jumping in a period of time shorter than that required for the emotional reaction to take place in the presence of the CS, we can no longer argue that jumping continues to be reinforced by anxiety reduction. Then, since no further reinforcement occurs, the jumping response ought gradually to extinguish.

The data described in Experiment 2 indicate that the animals *did* become less and less emotional during the ordinary extinction trials. However, there is substantial evidence to indicate that the conditioned emotional responses which were established early in training had not, in reality, been *extinguished* at the end of 200 trials of ordinary extinction procedure. It will be remembered that most of the dogs in Experiment 3 were profoundly

upset when introduction of the glass barrier first restrained them in the presence of the CS. Though, during ordinary extinction, with very short latencies of the instrumental act we see no emotional reactions, when the animals are later held in the presence of the CS by the glass barrier they demonstrate that the CS has *maintained its capacity to elicit anxiety*.

But if the emotional response has not been extinguished, what has happened to it? The effect of early reinforcements has shortened the latency of jumping to the point where, since the dog's jump removes the CS so quickly, the conditioned emotional reaction may not be elicited at all! This would save the emotional reaction from extinction, since it is no longer exercised. But, if the emotional reaction is not elicited, we are now confronted with the picture of a dog continuing to jump without further reinforcement!

The law of effect would account for the persistence of a *certain number* of jumping responses after the last reinforcement. The exact number would be a function of the amount of reinforcement received. The law of effect would then predict a lengthening of latency, leading toward extinction of the jumping response. But, *what will happen if the latency of jumping increases?* If, as we have hypothesized, the conditioned emotional reaction has not yet been extinguished, a longer latency will leave the dog in the presence of the CS for a time interval long enough for the emotional reaction to be elicited. (The glass-barrier procedure did, in fact, accomplish this.) The elicitation of the emotional response when a long latency occurs ought to lead to a jump which will in turn be followed immediately by anxiety reduction. This further reinforcement would again strengthen the bond between the CS and jumping. A cyclical mechanism of this sort, with the emotional reaction being elicited only rarely, but, when elicited, giving renewed strength to the CS-jumping bond, would account for the high resistance of the jumping response to the ordinary extinction procedure. The final extinction of jumping could begin only when the emotional reaction had been elicited often enough to be itself extinguished, or at a low strength.

We feel that the order of events which we have described is general to *all* learned avoidance responses. The picture is surprisingly akin to the clinical picture in compulsive neurosis. It contains the possibility of the organism "frightening itself" by remaining in the presence of the CS long enough for the CS to be effective, while no emotionality will be elicited with *short* latencies for *instrumental* acts. We feel that this interpretation can help us to circumvent the "dilemma of fear as a motivating force," which Eglash (1) feels is a stumbling block for anxiety-reduction theories.

This formulation suggests an interesting experiment. If the animal's jump can be delayed until *after* an emotional reaction has been elicited by the CS, and then the animal is allowed to jump, two opposing processes should go on. The habit strength of the jumping response should be augmented; but the elicitation of the emotional reaction without the presentation of the US (shock) should weaken the strength of the conditioned emotional response (anxiety). Of course, the best way to produce extinction of the emotional response would be to arrange the situation in such a way that an extremely intense emotional reaction takes place in the presence of the CS. This would be tantamount to a reinstatement of the original acquisition situation, and since the US is not presented a big decremental effect should occur. One of the main problems in extinguishing avoidance conditioning is, thus, keeping the animal in the presence of highly disturbing danger signals which are no longer followed by noxious stimulation. Usually, well-learned instrumental responses *prevent this from happening* and slow up the course of extinction. (It seems to us that the "reality-testing" procedures in psychotherapy are *partly* designed to face this problem.)

But there is a second principle, possibly just as important as the first. Extinction should be much easier to obtain if, in conjunction with the reality-testing procedure, the animal is not allowed to perform the instrumental avoidance response. Our glass-barrier procedure approximated this, since the CS was removed at the end of two minutes, and any anxiety reduction that took place was preceded in time by *responses other than jumping*. Thus, while only two out of nine of our

animals extinguished in 10 days of the glass-barrier procedure, on the basis of our argument we would be forced to predict that this procedure would, in relatively few additional trials, have led to extinction for all animals. There is some reason to believe that the reality-testing procedure must approximate the acquisition situation more closely than we were able to approximate it, in order for this prediction to materialize. The behavior of the dogs was at first puzzling. It will be remembered that they were initially upset by the glass-barrier procedure, but they quickly quieted down with successive trials, and later most dogs remained in a relatively relaxed state when the glass barrier was encountered. It was surprising to see the dogs jump quickly on trials on which the glass barrier was not present when they seemed to have "learned to relax" in the presence of the CS plus the glass barrier. We feel that they learned to discriminate the two conditions, with and without glass barrier, and that each condition controlled its own response. This raises an interesting problem: after traumatic avoidance learning has taken place, how do we fashion reality-testing procedures that are indiscriminally different from the acquisition situation, so that the instrumental act can be removed? (Presumably the psychotherapist, when forcing reality testing on a patient, has learned some way of doing this, or else the patient would relapse completely on leaving the therapist's office.) We are forced to predict that the glass-barrier procedure would be far more efficient than the ordinary procedure. Furthermore, an additional experiment, in which the animal is forced to delay jumping long enough to elicit the emotional response, but is then allowed to jump, should definitely be *less* efficient than the glass-barrier procedure we have already used.

While a two-process theory forces us into fairly definite predictions about the glass-barrier, reality-testing procedure, its application to the shock-extinction procedure is by no means clear-cut. A simple Thorndikian interpretation of the shock-extinction procedure might maintain that punishment of the jumping response ought to stamp it out. The data do not support such an interpretation. (Neither do the data presented by Gwinn (2), on the use of punishment in the extinction of a running response in rats.) Only 3 of 13 dogs met the extinction criterion during 10 days of the shock-extinction procedure. The other 10 persisted in jumping, and on the average their latencies shortened. Gwinn has reported that his rats, when shocked for running, at first increased their running speed, but with additional shock trials some of them gradually extinguished, and the others were discontinued without reaching an extinction criterion.

From our point of view, the failure of the shock procedure to bring about extinction in 10 of our animals might be explained with the help of an additional assumption. In the shock-extinction procedure, we would clearly expect a few jumps on the basis of previous reinforcement of jumping. However, during the initial jumps into shock, the CS is once again followed by the US, though it is true that the jumping response intervenes between the two events. We assume that the renewed pairing of CS and US, by contiguity principles, drastically strengthens the fear reaction to the CS. At the same time, the habit strength of the jumping response should be weakened through Thorndikian action of punishment. We would then have to claim that the increase in drive more than counterbalances the decrease in the habit strength of jumping. We observed that emotionality between trials increased greatly during shock-extinction trials. The raised anxiety level probably affects the "operant level" for jumping in a shuttlebox situation. This phenomenon might overlap with the fear reaction specifically aroused by presentation of the CS.

This is clearly a dialectical argument in its present form, capable of accounting for *any* observable results! But it would not be difficult to test the formulation. Any technique which first reduces the strength of the anxiety reaction ought to increase the efficacy of the shock-extinction procedure. This follows, since the hypothetical events crucial for extinction are the lowering of drive and the lowering of habit strength. If, at the beginning of the shock-extinction procedure, drive (anxiety) is low, then the Thorndikian action of punishment can presumably take place *before* the new series of shocks can build up the intensity of anxiety in the presence of the

CS. We can then predict that if the use of punishment, increasing the anxiety level, *follows* a reality-testing procedure that is continued until the conditioned emotional reaction has diminished, it will be much more effective than if the punishment procedure *precedes* reality-testing. If, however, because of previous procedures the anxiety level elicited by the CS is fairly high at the beginning of the shock-extinction procedure, the new shock series will drive the anxiety level to a point where the decremental, Thorndikian effects of punishment cannot be observed. Thus, we can predict that *if extinction does take place with shock-extinction procedure, it must do so within very few trials, or else be extremely difficult to obtain.*

Our data are not in disagreement with these predictions. Of 10 animals who received the shock-extinction treatment before the combined treatment, two failed to extinguish during the combined procedure. Of six animals who received the glass-barrier procedure before the combined procedure, all extinguished during the combined procedure. In addition, the animals in the latter group required fewer trials of the combined procedure, as well as fewer shocks, to meet the extinction criterion. Using the shock procedure alone, those three animals which did extinguish did so in less than a dozen trials, while those who did not extinguish were jumping with short latencies after 100 trials. A more crucial experiment would have employed the shock and glass-barrier procedure in sequence, counterbalancing for order of presentation, without introducing the combined procedure. Then the sequence of glass barrier followed by shock should lead to rapid extinction.

Our interpretation of the shock-extinction procedure might be further sustained by an experiment in which the punishing shock is of low intensity. Here, the increment to conditioned anxiety might be less than the decrement to habit strength. Extinction might then take place quickly if the low-punishment procedure were preceded by the glass-barrier procedure.

The results of Experiment 5, on the combined extinction procedure, are not at odds with our theoretical discussion. This procedure, as we would expect, was more effective when it was preceded by glass-barrier treatment than when preceded by the punishment procedure. The combined procedure was effective in 14 out of 16 cases. Its effectiveness probably derives from the fact that the reality-testing procedure tends to reduce the strength of conditioned emotional reactions. This reduction offsets the increase of anxiety produced by the punishment shocks, at a time when the habit strength of jumping is being decreased through punishment. The result is extinction of jumping.

While the data of our five extinction experiments are fairly well fitted by our theoretical formulation, we are left with many points of difficulty, as well as some circular reasoning. Despite such shortcomings, the fact that the argument suggests new, independent lines of attack on the problem of extinction of traumatic avoidance learning is encouraging.

On most points, our interpretation of traumatic avoidance learning is in essential agreement with both Miller (5, 6) and Mowrer (7, 8). We agree with them that the development of the acquired anxiety drive is essential to the establishment of avoidance conditioning. We differ with Mowrer in including skeletal responses among those emotional reactions which are susceptible to classical conditioning. We differ with Miller because we feel that a one-process, S-R reinforcement theory is inadequate. We do not believe that anxiety is reinforced through the law of effect, but rather through principles of stimulus contiguity. In this respect, we tend to be more sympathetic with the views of Schlosberg (10) and of Skinner (11, 12).

In carrying out our analysis of the rather minute details of behavior of dogs in a traumatic avoidance learning situation, we were faced with several inadequacies in current theory. Our own theoretical formulations have been, for the most part, improvised to account for our own findings. They and those of Miller, Mowrer, and Schlosberg are not mutually exclusive. At present, all of the theoretical alternatives supplement each other and help to order the data on avoidance conditioning.

SUMMARY

We have described a series of five experiments in which several different procedures

were used to bring about the extinction of traumatically induced avoidance responses in dogs. These were the main findings:

1. With an ordinary extinction procedure, cessation of responding was extremely difficult to obtain.

2. A glass-barrier, reality-testing procedure was moderately effective. In this procedure the animal was detained in the presence of the danger signal without being shocked and the instrumental avoidance response was not allowed to occur.

3. A shock-extinction or punishment procedure was approximately as effective as the glass-barrier procedure. In this procedure the animal was punished for making the instrumental response.

4. A combination of the reality-testing and punishment-for-responding procedures was very effective in producing extinction. This combination procedure was more effective when preceded by reality testing than it was when preceded by the punishment technique.

These findings were interpreted from the point of view of a modified two-process learning theory similar to Mowrer's (8).

REFERENCES

1. Eglash, A. The dilemma of fear as a motivating force. *Psychol. Rev.*, 1952, **59**, 376–379.
2. Gwinn, G. T. The effects of punishment on acts motivated by fear. *J. exp. Psychol.*, 1949, **39**, 260–269.
3. Maier, N. R. F., & Klee, J. B. Studies of abnormal behavior in the rat. XII. The pattern of punishment and its relation to abnormal fixations. *J. exp. Psychol.*, 1943, **32**, 377–398.
4. Maier, N. R. F. Studies of abnormal behavior in the rat. XVII. Guidance versus trial and error in the alternation of habits and fixations. *J. Psychol.*, 1945, **19**, 133–163.
5. Miller, N. E. Studies of fear as an acquirable drive: I. Fear as motivation and fear-reduction as reinforcement in the learning of new responses. *J. exp. Psychol.*, 1948, **38**, 89–101.
6. Miller, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
7. Mowrer, O. H. Anxiety-reduction and learning. *J. exp. Psychol.*, 1940, **27**, 497–516.
8. Mowrer, O. H. On the dual nature of learning—A reinterpretation of "conditioning" and "problem solving." *Harv. educ. Rev.*, 1947, Spring, 102–148.
9. Mowrer, O. H. Learning theory and the neurotic paradox. *Amer. J. Orthopsychiat.*, 1948, **18**, 571–609.
10. Schlosberg, H. The relationship between success and the laws of conditioning. *Psychol. Rev.*, 1937, **44**, 379–394.
11. Skinner, B. F. Two types of conditioned reflex and a pseudo-type. *J. gen. Psychol.*, 1935, **12**, 66–77.
12. Skinner, B. F. *The behavior of organisms*. New York: Appleton-Century-Crofts, 1938.
13. Solomon, R. L., & Wynne, L. C. Traumatic avoidance learning: acquisition in normal dogs. *Psychol. Monogr.*, 1953, **67**, No. 4 (Whole No. 354).

Received November 21, 1952. Prior publication.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

CASE REPORT

REFLECTIONS BEFORE THE GALLOWS: THE LAST TESTAMENT OF JEROME TITUS

SUBMITTED BY PERCY BLACK [1]
University of New Brunswick

At the closing of a long life it is not uncommon for a person to feel an inner urge to look back upon himself for a final tally. Often he experiences little fear of the last moment. He becomes resigned to the inevitable. Not man, but powers beyond any man's control subdue the desire to live on. Nothing one has done, nothing one can do, can alter this last invincible sweep. One merely lies back and waits.

When one is young, however, surrender comes hard. The game is not yet played out. Death is still a distant improbable. When death leaps backward from a comfortable distance in time to confront one suddenly with its imminent certitude, youth cannot lie back and wait; it fights, it relentlessly examines itself, it alternately fears and courageously faces death. This is especially apparent when it is men, not God or nature, who thrust upon one the starkness of an imposed cessation of being. The last testament of Jerome Titus offers such an inglimpse into the workings of a clichéd mind under the conditions of maximum earnestness and strain.

To appreciate more fully the context of the case, the following details may be in order.[2]

Jerome and his brother, Horace, Negroes, aged 23 and 22 respectively, lived together in Jerome's flat, where also lived Jerome's wife and their two children. Horace was single. One night these brothers decided "to look for girls and have a good time." Being without jobs and having no money for liquor and *filles de joies,* they determined to rob someone—any likely prospect on the streets of River Town. "We was going to hit a fellow . . . any fellow driving a taxi," Jerome later testified. But Jerome, although having the responsibilities of a husband and a father of two children, followed with reluctance the urgings of his younger and less responsible but more dominant brother.

Fortified with liquor from their meager savings, Jerome telephoned several taxi stands for a taxi to come out to where they were, near the outskirts of the city. After several unsucessful calls, one taxi finally arrived. Jerome held a hammer in readiness to strike but when he recognized the driver, Corey Willows, as one of his former employers, who had treated him well, he "lost his nerve" and couldn't wield it. Horace grew impatient and enjoined his brother contemptuously to "go through with it." Meanwhile the chauffeur was asked in the usual manner to drive them to the haunt of a well-known bootlegger.[3] This, Willows did without question.

Jerome then left Horace and the driver chatting in the car and went into the bootlegger's lean-to. On his way, Jerome, fearing that Horace might later influence him to carry out something "too risky," or to execute his desire by himself, secretly disposed of the hammer at the rear of the car. After some minutes, Jerome returned to find that Horace had discovered the hammer and had struck Willows unconscious. The deed consummated, they pilfered several hundred dollars from the bleeding driver, squeezed him into the trunk compartment of his own taxi, and rode off to make merry in an adjoining city.

Experiencing guilt at the enormity of their crime, and bewildered by the uncertainty of Willows' condition, they hoped to rid them-

[1] I am indebted to a former member of the Salvation Army who was at the side of Titus to the last, and through whose generosity the document came to my attention. Names and places are disguised here to avoid embarrassment to all persons mentioned in the document.

[2] These details I have collected from the files of newspaper reports on the apprehension and trial of the two brothers, as well as from persons connected with the case.

[3] River Town is a "dry" city where, as usual, spirituous beverages may be illegally obtained "if you know the right people."

selves of the apparently lifeless form by throwing it into the sea. Apparently this plan was abandoned, possibly because there were too many onlookers. Failing this, they contrived to remove the blood stains from the car seats. Jerome asked his mistress of the night for water and rags because, as he said, "a couple of my pals got sick in the car."

Their spree over, they drove back to River Town and parked the car on a deserted road where the body was soon discovered. Later, on the basis of the autopsy the coroner concluded that Willows had not been dead when placed in the trunk, that he bled to death, and that his fingernails were broken, evidently in an attempt to claw his way out.

After some investigation the local authorities tracked down the two brothers and apprehended them on a charge of petty thievery, previously committed individually, to which the brothers separately pleaded guilty. Earlier Jerome had tried to burn the murder weapon and various articles belonging to the deceased, but fingerprints and other evidence led to their formal conviction in connection with the murder.

During the first days of the trial both brothers parried questions and pretended innocence. Later Jerome, apparently recognizing the futility of bucking the mounting evidence against them, confessed all, hoping in this way to escape the supreme penalty. The court, however, convicted them both—Horace for first degree murder, and Jerome as an "accomplice and a party to the offence." They were sentenced to hang back to back on the same day.

Awaiting the day, both brothers adopted religious affiliations, having previously been churchless—Jerome joined the Salvation Army, and Horace, the Catholic Church. Since neither of them was a newcomer to crime and dishonesty, one can only surmise the genuineness of the motives behind their new credos. Judging by the intensity of religious expression in Jemore's testament, by his repeated exhortations to the world to turn to God, by his forgiving of hurts done him by others, by his penitence for his former unbelief, and by his dreams and visions of Christ, we may safely conjecture that for him, at least, conversion was sincere, and not just a stratagem to conscript clerical and public compassion for his craving to avoid the gallows. Certainly it functioned for him, as he tells us, to ease him through the strain of waiting for the approaching inevitable. Though the trial was eminently "fair," as Jerome himself testifies, right to the end he wished ardently for a last-minute reprieve. His simple faith in God's material reciprocation to man's prayers probably served to intensify this wish, especially after his awful, yet reassuring, experience of the presence of God's spirit in his cell.

Jerome Titus tells his story with the simple earnestness of a defeated child pleading forgiveness before an outraged mother. It is an emotional outpouring with cathartic intensity, perhaps the only recourse left to a trapped, guilt-ridden mind trying to keep a semblance of sanity before the unremitting backwash of death's crush.

June-6-1949[4]
By Jerome Titus,
2619 Dix St.,
River Town

i[5] die at midnight on July the, 27, of 1949, on Wednesday morning early, when you read this i most likely will be dead, but don't be alarmed at hearing from a dead man, for now, as i begin this story, i'm very much alive. it June the 6 Monday, 1949, midnight Wednesday July the -27-1949. i am scheduled to die for murder. Sitting here in my cell in Court County Jail, in the city of River Town, i've been doing lots of thinking, some of my thoughts are a warning to criminals, and my wife and my to children a boy and girl and to all sinners of this world which God our Heavenly Father made for us, and this story really is just part of my life but its the truth, the real story, i feel, lies in the fact that i don,t mind talking about dying,

[4] The reader will note that about one half of the document was written during this day, and almost the whole of the rest on the following day. During the interval between June 7 and the day of his execution, Titus discontinued his testament waiting "to see what happens in the next six weeks to come." On the day before the hanging he completed his last lines.

[5] Only a minimum amount of editing has been resorted to, involving for the most part the leaving of spaces at the end of thoughts. Otherwise, as far as possible, the original pencil-written document is reproduced here almost identically as Titus bequeathed it.

i'm just 23 years old but i'm ready to go, you see why, if my number were up this very minute, i'd be ready to meet God, i'm really happy, just this week i had a dream that i'll carry with me to my grave, i was on my way to heaven, Jesus Christ the son of God was with me but i was taking six steps to his two Jesus ask me why i was going so fast i told him i was eager to get there, then i was there, surrounded by God angels, some folks might think that's strange talk from a man who came to death cell as a murderer but that's just the way i feel, you'll understand better when i tell you how i met God early one night, but first, take a glance at my past life, that i had when i was a kid, you see i never had a mother or a father to love me and bring me up like other kids, you see i brought my self up into this world, when i was 13 years of age i was out working for my own living and making $10 a month, this happen when i left the Children home, and i only had grade three schooling, so when i was 17 years old i joined the Canadian Active Army, i was discharged with a honourable discharged A-10-29-10 in July the 20,-1944, and then i joine the Merchant Navy and was honorable discharged from there, you see i had strugled thrugh life the hard way and was kick and beat around, and then my crime life started five years ago in Montreal Que, i was sent up for stealing a car and $4,000, dollor woreth of furs, you see i got with a bad crowd and then i got in to truble, well i did my time and came out i was only a kid a the time, my age was 19 years old, and i did one month and two month in Montreal prison again, so i said i would go straight, so i left Montreal and came to River Town to start to lead a good life and settle down and be a man you see i am all right when i am left alone. Well i got married to a good women now i been married two years and five month and i am the father of a boy and a girl, one of my children is 17 months old and the other 6 months old, i was happy man untill my brother came out of the pen, and neither one of us had any money so my brother and i went out to rob a man and get some money but my brother hit the man a little too hard with the hammer, so it killed him, so my brother and i was tried for first degree murder and was convicted and sentenced to hang on July 27-1949 on Wednesday night at midnight, i am tell the world you cant beat the law so dont try it, because i did and now i am faceing the gallows for my life but i will never lose fath in Jesus Christ, i will all way love my Saviour and i will praise his name on the gallows and give thanks to Him, so after i had come out of the prison from doing my time i had led a good life and made an honest liveing, untill that night back in January the seven 1949, when my brother and i went out for money and a man was killed in Rockhaven out side of the city, and no body know the sorrow in my heart and how sorry i feel inside except God, for what my brother and i have done, we had taken a mans life for a hundred and eighty dollars, so i say he was a good and kind hearted man and he was a married too with three children and my heart is in great sympathy and great sorrow, and i have sent word to his wife to ask her to for give me and that i am sorry to be micked up in the death of her husband and my friend,[6] so i am writing this to the whole world that breaking the laws dont pay you cant beat the laws of this earth or the laws of God so live honest and work with your hands for a honest living and belive in Jesus Christ and read your Bible and you will get along in this world and in the next world too, not long after i was placed behind the bars last January the-11-1949 before Mayor Love of the Salvation Army came to see me, and told me about the great word of God our Havenly Father and His love for me, and eatch and every one of us sinners, come to God and kneel before him and confess our sins with thy mouth and ask forgiven and he will for give us and save us from hell, so the day that my [unfinished thought; space left by Titus] Mayor Love came i was playing cards on my bed and reading all kinds of filthy books a funnys books too, and Mr, Marshall and his wife pleadeth with me to except Christ as my personal Saviour i laughed at them all, i had said to my self what can God do for me a man charged with murder and had sin againest God why i dont believe there's is a God i boasted, but i found out better for my self, because i have read

[6] Titus had formerly worked for the murdered man and had been on good terms with him.

God word and have found Jesus Christ, and i believe with all my heart that i am forgiven of my sins and am saved, i have given up every thing and all of my bad habbet for Jesus sake, for [i.e., four] month ago, and he love me and i love him always, you see i woke up one night sweating although the cell was cool i new that some body was in my cell and that was the spirit of God he spoke to me and said to me why dont you beleive in me is it because you cant see me you believe there is a king and you have not seen him so now why dont you believe in me, well i come to you to night in your dreams and he did and then i knew that God was real and mercyful to sinners, no body had seen me climb from my bed and fall on my knees and cried like a baby to God, i don't remember just what i told God that night on my knees but i asked him to be merciful to me an evil sinner, he saved me that night i know i've believed on his son Jesus ever since. well i hated my brother for coming back in to my life and ruining it and all so my wife and children life to, but God had saved me from sins so why should i hate my brother or any body else God had taught me to love every body so i do, later while i was in here the word got around that i was a Christian and that i was putting on a Christian act trying to get out of going to the gallows but believe me sinners i wasn't putting a act i mean every thing what i said about the love i have in my heart for Jesus Christ and the love that Jesus has for us all, sure when the judge gave me the death sentence it jolted me i all most drop dead, but i haven't lost faith in God because if God had not been with me all the way i could not have stand up to it all by myself, or face the people i was ashamed of my self, because you see i was a well like guy in this city onetime, but i dont blame the people for being down on me, i blame my self for what is happen to me, i no right from wrong, and i should of not let my self be lead by the nose by any body, and i had a fere trial, so i blame my self Jerome Titus for my wrongdoing so you see i know God will go with me, so you see i'm really not afriad but before i die i want to leave one last message for other young people, start serving the Saviour now our Lord Jesus Christ while you're young, grow up this way and it'll keep you straight, once crime get a hold of you it,s hard to stop just like the habits of smokeing and drinking if they once get ahold of you, you can,t quit, so please take my advice, the wages of sin is death. But the gift of God is eternal life through Jesus Christ our Lord, (Rom, 6–23) i found out it,s true and for me to live is Christ and to die is gain the Bible is true and God word to us and His laws, lessen at God love for us and the comfort to us, the Twenty-Third Psalm: 23:4.

Yea though i walk through the vally of the shadow of death, i will fear no evil. For Thou art with me, Thy rod and Thy staff they comfort me, nowthere are no sweeter words in the Bible than them God letting us know that he is beside us even unto death if we love Him, the way he love us, and now i pray for every body and tell God i love every body, then i prayed for people i made suffer, i pray for my mother and father and sister who i had hated because the way they treated me i was the black sheep of our family but now i love them all, and i ask the Lord to bless them and for give them of there sins, so now i am praying for a life sentence in prison not because i am afraid or i am not ready to die because i am, but it is on account of my wife a our to children i have brought anought of disgrace to them already and to know that thay will grow up here and some body say to them your father was hanged for murder on the gallows it woulden sound so good would it, that the reason i dont want to go to the gallows is on account of our children, you see i never had a chance at life like any bodys else, and i would like to get that chanch to work for God someday so i put everything in God hands now, what every is God will and wish it will happen so my faith in our Lord Jesus Christ is the strongest in the whole world, and i no he will answer my prayers,

The jail that i am waiting here in is very clean and very good foods, and Mr James is sheriff of the jail and county, and Mr, Mathew is sedond sheriff of the jail and run it and keep it in good order, and let me tell the whole world it the best jail i ever been in to, and we have three guards on here to watch us and they are very good to us, and all so the sheriff is very good him and Mr. Mathew

is the best in the world, and remember dont take your own life because you are lost forever in hell, so pray for me and that i will have good luck, and may God bless you all, Amen.

Now this is Tuesday June the 7–1949 and i am 'wrighting some more as i have told the world about my past life, if i had of belive in Jesus Christ and trust God i would not be here to day faceing the gallow for my mistake, iwould be home looking after my wife and children, so sinners now is the time to put your whole life and trust in God now will you do that this very day please, dont give part of your life to the Satan, and the other to God because God is not satfide with half of your love he wont all of it because God give us all of His love no matter what we are God love us all, and just look here how God showed his love for us all, for God so loved the world, that He gave His only begotten Son, you see i ask you to think about this, John–3–16, now God tell us, that who so ever believeth in his Son should not perish, and you see sinners God has given us his word and all we got to do is except Christ in to our heart as our personal Saviour that all God ask from us so why not do that to day, read your Bible and i will go so far as to say that you will get all the foundation and happyness you wont to stand on and God wont let you down or break His word to us, if we believe in His Son Jesus Christ, God has promist us a lot all ready, but look what he still is willing to give us if we believe in His Son Jesus, but have ever lasting life isen that great news to here, just to set down and think that there is a way to live for ever in heaven with Christ the Son of God. it never to late to accept Christ, dont wait like i did before you accept Jesus Christ it might be to late, do it now while you have a chance because you might die to day or to night or you might walk out in front of a car and get kill and if you are a sinner and you have not except Christ at this point, you are lost in hell fires and there no hope for you, you dont have to believe me what i am writing, all you got to do is read God word and that the Bible, and that what i did to be saved from hell fires, i read my Bible and found Jesus Christ and then i confessed to God my sins and ask for given and he did for give me, well you might say to some body how does he know that God has for given him of his sins, well all i can do is say i believe in my heart that God has for given me because i feel it in side of me, and let me tell you that there is no body on this earth that can take you to heaven with them or can for give you of your sins, they can pray for you and help to find our Saviour Jesus Christ and that all they can do for you, i my self listen to what was told to me about Christ, and dont for get the only thing that can save you from hell is Jesus Christ blood clened us from all sins, and our sins are wash away and never more rember by Christ, isn't that good news from God him self, and we dont have to pay out our money for it, it given to us for nothing, all we got to do is accept Christ his son, now you no about me i found my Lord and Saviour in jail and i am happy about it to because i found Jesus Christ before it was to late, you see what Christ say here, blessed are they which do hunger and thirst after righteousness for they shall be filled, St, Mat–5–6, you see if you are seeking Jesus love he will fill you and come to you, and if you are thirsty after Christ joy and peace and His salvation and comfort Christ will give it to you if you accept Him in your heart, and you will never thirst again in this world or the next world, and Jesus tell us if we are pure in our hearts we shall see God and you will find it in, St, Mat,–5–8 we shouldn't need any more then that of Jesus Christ His word, that was enough for me to believe in Him so let it be enough for you how about it what do you say, confess on your knees and open your heart to a new light and come to Jesus Christ and ask Him to for give you of your sins and have mercy on your soul and do it now and see if your not a new person and feel the glory of God shine in you try it now will you and i mean you, now Christ tell us to confess Him before man, and he will confess us also before his Father witch is in heaven, so what does He mean well i will try and tell you, well He mean for us to tell what a wonderful Saviour He is and that He can save our souls, and to tell every body about His love for us, and remember this Christ died on the cross for our sake He gave His life for the world on His own free will i my self was a very deep sinner

and i am only saved by the grace of God, you no Christ tell us to come unto Him that are labour and heavy laden and He will give us rest so you can see He a wonderful Saviour to have on our side, except Him now, St, Mat, 11–28 and dont for get money cant buy your way into heaven that one thing money cant by is the love of God for us, now Jesus tell us that a rich man shall hardly enter into the kingdom of heaven, St, Mat, 21–23, you see what Christ thinks about money so for get about money and except Jesus Christ as your personal saviour now as you are His love is far greater then all the money in the whole world, well you ask me how do i know because he showed me His love and the faith i have in him, He showed me how to help others who need it and how to be honest and live right i was a little late in finding out but not to late, why do we go around hurting other people that what i am thinking to night why do we do these things, and the only way that you can over come this is to except the Lord Jesus and lean on him and let Him take your troubles on His shoulders He strong and firm anought to stand it all we are to weak, i am writing this for a lesson to some poor sinner that needs Christ, you see what happen to me dont you, well dont wait tell you get in jail or faceing the gallow before you find your Saviour Jesus Christ, do it now to day so that my life story and how i found Jesus Christ, now as i set here in the County Court Jail in the city of River Town at 24 River Street,[7] i am going to tell you a little about this jail, well the first i am going to tell you about how clean Mr. Mathew the second sheriff keeps it, he all so runs the jail, and let me tell you about the good food here you get stews and beans and soups and jam and cheese and milk tea meats, and i am allowed to see my wife and children three times a week, and the food is very good in here i think so any way, and i dont care if you go all over Canada this is the best jail i have been in for all around cleaness and food and all around sports,

Mr. Mathew the second sheriff is the greatest man in this world to look after a prisoner, and the first sheriff Mr James is the best man in the world, they both is good men they let me have a radio in my cell and a clock, and any thing that my wife brougt me in i was allowed to have it, you no they dont half to be so good to my brother and i who was convicted of murder no they could shut us away in a dark condemn cell and leave us there untill meal times, no but they give us a big comfortable cell, and i myself never knew that there were such good people in the world as Mr, James the first sheriff and Mr, Mathew the second sheriff of this jail, God bless them always. and we have Mr, Tracey a guard here watching us and he a very good man, and will do any thing in his power that he aloud to do for us, and we have two other guards here to, but they all take there order from Mr. Mathew the second sheriff, and now i have one month and twenty two days left[8] before the last hour comes it funny isen it that a man life is cut so short on this earth and he know he going to die on the gallow isen it, but the funny thing about it is to show you the power Jesus Christ got in this world and the next world to, you see Christ has come into my life and made me a new man and made me not afraid to die, but i am hoping and praying to God that He will give me another chance to prove to the world that i can be a better man and a good father and husband to my wife and children, that all i pray for now, because you see i am not a murerer or never thought any thing about killing any body in my liff but the evidence was against me because i told the truth about the whole crime and that why i am here today by my own mouth, but after i walk that last mile i want the whole world to know this that i did not know that my brother would hit Corey Willows, i did not have any thing to do with the murder of Willows or the planning of it, untill i came back from Jacks, and seen Willows, laying over the seat of his car and then i help to take the money after he had been hit by my brother when i was down to Jacks, i am innocent of the murder of Corey Willows,

[7] The reader will note that this address is different from the one given at the beginning of the testament. It is possible that the first address was that of his residence, but I have not been able to establish this as fact.

[8] Upon checking the correctness of his calculation, we find a total of 50 days between June 7 and July 27, i.e., one month of 28 days plus 22 days more. Titus shows in this and other references that he was well oriented with respect to time, a good indication of his mental balance.

i would say that i am accessory after the fact i am guilty of that part, and i would like the people of River Town to believe that after i am gone, and now i am praying for my family, that God will watch over them and give me another chanch to live in this world, some place else with my family and bring my children up the right way in life, and ever since i been in here i been praying for Mrs Willows and her children and may God let her forgive me for the part i had in the death of her husband and may God bless her and her children always and guide them through life, i will right the rest of this story later, i am going to see what happens in the next six weeks to come, i haven't given up faith yet God will answer my prayers i no He will one way or the other way, and may God bless every body in the whole world, if i go to prison for life witch i am hopeing and praying that God will answer my prayers, they say if you believe in Christ and ask any thing in His name it shall be given unto you, so i am praying night and day in the name of Jesus Christ, in the name of our Lord Jesus Christ may be glorified in you and ye in Him, according to the grace of our God and the Lord Jesus Christ, and may the grace of our Lord Jesus Christ, be with you for ever and ever A,men praisf His name Jesus A,men A,men. July-26-1949, now this is July the 26 and it is about 230 in the morning, and the just put two new guard on hear, and i was watching the men build the gallows to day for me and my brother, so after 12 o,clock to night it will be July the 27 and we are going to die then, unless it God will for us to live in this world again, but i am ready to meet Jesus Christ our Lord and Saviour i am not afried to die becuse Jesus tell us in His holy word, for what shall it profit a man if he shall gain the whole world and lose his own soul, St, Mark-8-36, I am talking to all sinners your soul are the one you got to look after and not your body becuse your body is made of dirt of the earth and we shall go back to the earth where God took us from. But our soul live and that where all our sins lays and some day we all got to answer in front of our Lord Jusus Christ, for ours sins, and the only key to the gates of heaven is Jesus Christ our Lord, and my brother Horace is a Catholic and now it 8.30 in the morning of the 26 of July-1949 and they are holding Holy Mass in my brother cell now, but i am a Prostant i blong to the Salvation Army the army of God a soldier of Christ i know in my heart that the only way any body can have for givens of his sins is through Jesus Christ and your soul saved, and the only way you can get to heaven is by and through Jesus becouse He tell us so in His holy word i am the way & truth & life, Jesus our Lord & Saviour say that, now i am going to finnish my life story, no i just got word that i was going to to hang to night for shore well i know i will past from death un to life i am going to Jesus Christ our Lord and Saviour praise his name always. Now it is eleven o'clock things are coming to an end, and I am ready to go, because I know that Jesus Christ our Saviour has forgiven me of all my sins. Praise His precious name.

At[9] 7.30 Major and Mrs. Love came for a service. They brought the Band and were accompanied by Brother Isaacs and Brother Willy. In the Band were Bandmaster Thorne, Mrs. Thorne, Kenneth Hoyt, Phillip Randall, Ben Avery and Lorna Mackintosh. They played songs I had selected for them. I gave my testimony and listened to the testimony of others. The service was brought to a close with the song, "God be with you till we meet again." Major Love, Brothers Isaacs and Willy are now with me waiting for two o'clock when I am going to be with Jesus.

My wife attended the service. God bless her!

July 26th, 1949 by (Sgd) Jerome Titus

Titus wrote the following letter to one of his guards one day before his execution. I am grateful to this guard for granting permission to use the letter for publication.

July 26/49

Dear George sir

I wont to thank you from the bottom of my heart for all you have done to make me happy in my stay hear, my Life is coming to a end now, and I wont to make every thing write with every body, so if I said any thing to you or done any thing to hurt your feelings I ask for your forgiven. I have made my

[9] These final lines were written by someone else. Apparently Titus dictated them to a representative of the Salvation Army.

Peace with God, so if it God will I am Ready to go and I wrote a lot of letters out and told People how Good you Guards were to me, may God Bless you and you family

Your friend Jerome Titus
Cell -1-

COMMENT

To understand the intricacies of men's motivations requires more than a brief autobiography. The task, however, of providing such understanding can be immeasurably lightened when we have such self-reflections as in the Titus document. We ought not expect the insights of social and psychological science to speak to us through the blunt constructions of unlettered men. But even those who are not highly educated think and feel, and when they can express themselves on paper, howbeit muddled, they present us with a precious tangential entrance into their phenomenological dynamics. For a proper study of the social personality, W. I. Thomas (2) long ago recognized and advocated the essentialness of the subjective elucidation of the "personal life-record" for sociology, and similarly, G. W. Allport (1) has forcefully pressed for the riches of this approach for psychology. Only recently have there been signs of a quickening of interest in this medium, if we may assume R. W. White's (3) biographical analyses to be a brilliant earnest of things to come.

The document of Jerome Titus, though brief and inexplicit on many points that we should have liked him to tell us, invites critical study, nevertheless, as a source of data about mind *in extremis*. True, he himself was not aware of many of the conditions that impelled him to act. He could not tell us, for instance, why his brother, younger than himself—indeed dependent on him for board and shelter—had so much power to direct his behavior against his will. Yet, though himself greatly responsible for his plight, Jerome leaps above this most profound of his life crises to forgive his brother and all others who had made him suffer. Perhaps the approach of death tends to issue in the altruization of one's motives. Or perhaps Jerome was merely mouthing the dictates of an adopted creed. Whatever the verdict from a study of this impassioned document, it can hardly be doubted that were he to have gotten a reprieve and released into life a free man, Jerome Titus would never again have been the same.

There is cause to wonder what society loses by consigning such budding altruists to the care of the hangman. In cases such as the present it does not seem unduly far-fetched to guess that recidivism would be quite unlikely. But our morality in these cases, based on the precept of "an eye for an eye," is perhaps too powerfully ingrained in the texture of our civilization to expect any immediate lifting away of this ethical absolute. Meanwhile, Jerome Titus passes into social immortality by way of his testament. For he has given us an unusual glimpse into the self-sight of a man forced by society at a young age to grip his life in tortured reflection during the few remaining weeks of an unhappy life.

REFERENCES

1. ALLPORT, G. W. The use of personal documents in psychological science. *Soc. Sci. Res. Counc., Bull.*, 1949, No. 49.
2. VOLKART, E. H. (Ed.) *Social behavior and personality: contributions of W. I. Thomas to theory and social research.* New York: Social Science Research Council, 1951.
3. WHITE, R. W. *Lives in progress.* New York: Dryden, 1952.

Received March 14, 1952.

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

CRITIQUES AND NOTES

THE ETHNIC PREJUDICES OF WHITE AND NEGRO COLLEGE STUDENTS

J. STANLEY GRAY AND ANTHONY H. THOMPSON
University of Georgia

BEGINNING with Bogardus in 1925,[1] numerous studies have been made of "social distance" and the factors which affect these prejudices. Fifteen such articles were listed in *Psychological Abstracts* for 1951 alone.

TABLE 1

EFFECT OF ACQUAINTANCE WITH ETHNOLOGICAL GROUPS ON ATTITUDE TOWARD THOSE GROUPS

ETHNOLOGICAL GROUP	WHITES *		NEGROES *	
	TOTAL	ACQUAINTED	TOTAL	ACQUAINTED
Canadian	7.2	7.3	4.5	5.0
Mexican	4.6	5.5	4.6	4.6
Cuban	5.1	6.0	4.8	6.3
Englishman	7.0	7.0	4.7	5.1
Frenchman	6.5	6.8	4.5	5.9
German	6.3	6.6	3.0	4.8
Spaniard	5.5	6.4	4.1	5.3
Italian	5.6	6.0	4.2	4.7
Greek	5.4	5.0	3.9	5.4
Jew	5.7	6.0	4.5	5.3
Chinese	4.8	5.7	3.9	4.4
Japanese	4.5	5.0	3.1	4.6
Filipino	4.9	5.2	4.6	6.0
Total	5.6	6.0	4.1	5.1

* All scores given are means.

[1] BOGARDUS, E. S. *Immigration and race attitudes.* Boston: Heath, 1928.

This study is a comparison of 400 white undergraduate college students in the University of Georgia with 300 Negro undergraduate students in three state-supported Negro colleges in Georgia, using a modification of the Bogardus social distance scale. The students were asked to rate 24 ethnological groups on an 8-point scale ranging from "I would accept in my family by marriage" (value of 8) down to "Should be excluded from the country" (value of 1). An accompanying questionnaire obtained such information as age, education, religion, army service, parents' education, parents' income, and degree of acquaintance with each ethnological group. The rating sheet was not signed.

The average ratings for each ethnological group by both whites and Negroes are shown in Fig. 1. Note that the Negro students rate all groups except their own lower than do white students. The mean social distance rating (or prejudice score for 24 ethnic groups) by white students was 5.4, *SD* 1.07; that of Negro students was 4.1, *SD* 1.02. (The critical ratio of the difference in these means was 16.25.) The two distributions of frequencies of prejudice ratings are shown in Fig. 2.

It was found that acquaintance with at least five individuals of an ethnological group raised the social distance ratings of both whites and Negroes

TABLE 2

FACTORS AND THEIR EFFECT ON STUDENT ATTITUDES TOWARD ETHNOLOGICAL GROUPS

FACTOR	WHITE					NEGRO				
	N	MEAN	*SD*	M_1-M_2	*CR*	*N*	MEAN	*SD*	M_1-M_2	*CR*
Age										
17 and under	30	4.8	1.4	.8	1.78	23	3.8	.97	.7	2.8
25 and over	26	5.6	1.7			37	4.5	.95		
Education										
Freshmen	76	4.8	1.3	1.3	5.6	45	4.0	1.1	.6	2.6
Seniors	60	6.1	1.2			38	4.6	1.0		
Parents' Educ.										
H.S. graduate	292	5.4	1.4			34	4.0	1.16	.8	2.8
Col. graduate	95	5.4	1.4			24	4.8	.9		

for that group. Table 1 shows the ratings of ethnic groups by white and Negro students who were acquainted with at least five individuals in those groups compared with total ratings. With one exception (Mexican) the ratings of all groups raised with acquaintance.

Other factors which were found to affect social distance ratings were age, education, and the education of Negro parents. Freshmen were less liberal than seniors; older students were less liberal than younger students; and Negroes whose parents were college graduates were more liberal than those whose parents had but high school education. Religion and income of parents did not affect the degree of liberality. These data are shown in Table 2.

To check the reliability of the data, the same social distance scale was given to 100 high school

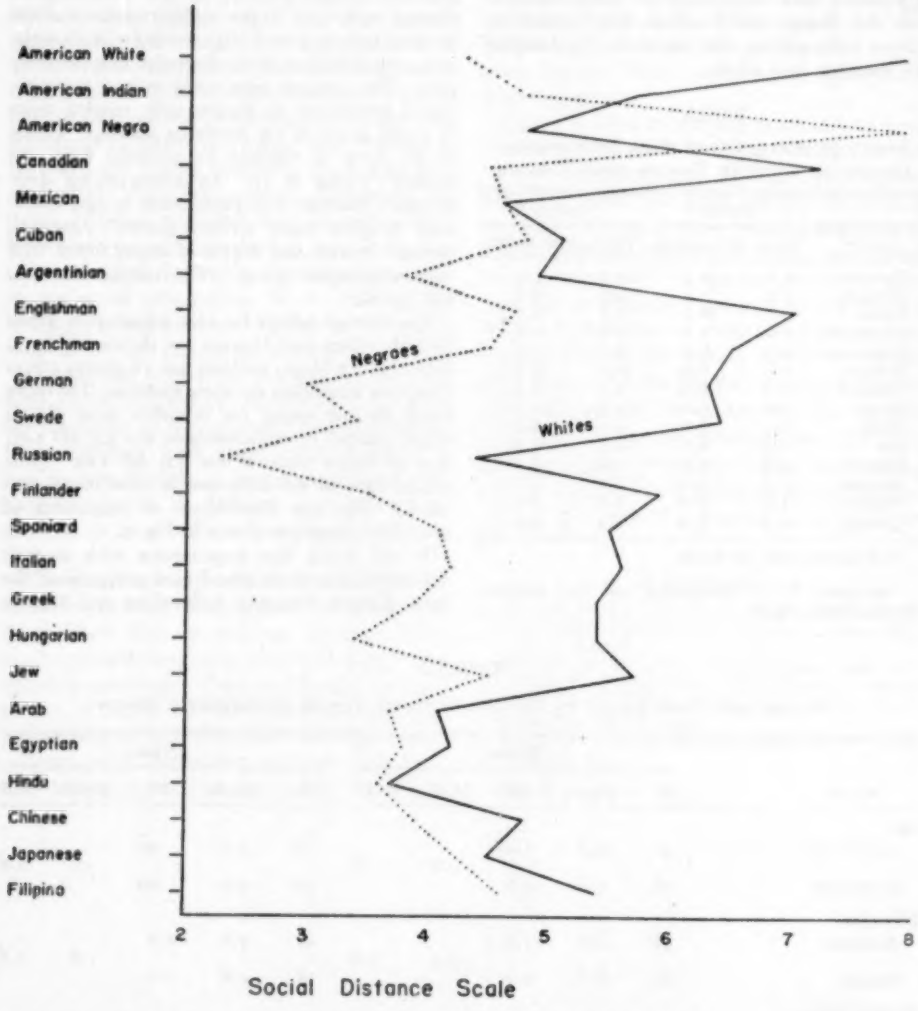

FIG. 1. SOCIAL DISTANCE RATINGS OF 24 ETHNOLOGICAL GROUPS BY 400 WHITE COLLEGE STUDENTS AND 300 NEGRO COLLEGE STUDENTS

TABLE 3

MEAN SOCIAL DISTANCE RATINGS OF 24 ETHNOLOGICAL GROUPS BY 100 HIGH SCHOOL STUDENTS AND 100 RANDOMLY SELECTED ADULTS

	WHITES		NEGROES	
	MEAN	*SD*	MEAN	*SD*
High school students	5.5	1.3	3.7	.75
Adults	4.9	1.3	3.8	.57

students (50 whites and 50 Negroes) and 100 randomly selected adults not connected with school work (50 whites and 50 Negroes). In both instances, as indicated in Table 3, the whites were more liberal than the Negroes, and acquaintanceship with nationality groups raised the social distance score for those groups. Both these check studies indicate that the data of the major study are accurate.

Received April 12, 1952.

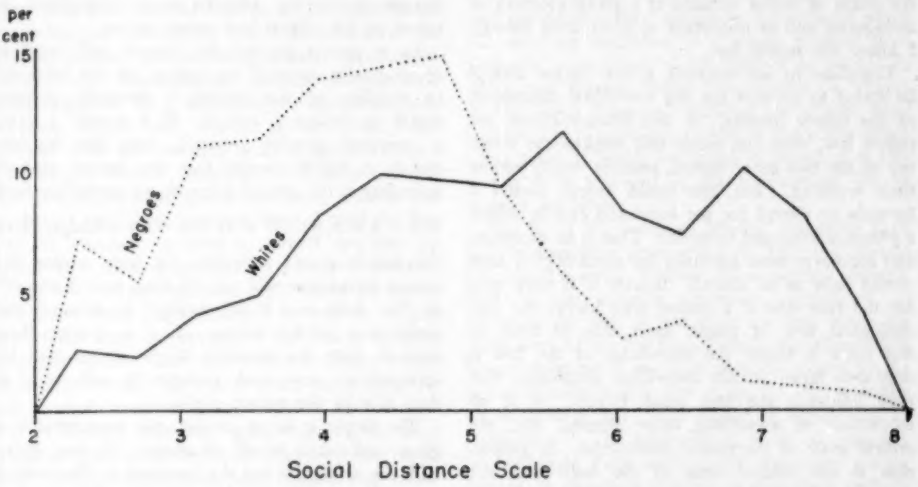

FIG. 2. DISTRIBUTION OF MEAN SOCIAL DISTANCE SCORES FOR WHITE AND NEGRO COLLEGE STUDENTS

THE JOURNAL OF ABNORMAL AND SOCIAL PSYCHOLOGY
Vol. 48, No. 2, 1953

THE BASIC LAW OF RUMOR

A. CHORUS
University of Leiden

ACCORDING to Allport and Postman (1), the two essential conditions for the transmission of rumor are importance and ambiguity. Roughly, importance stands for the emotional factor and ambiguity for the cognitive factor in rumor-spreading. Both factors are related to rumor transmission in a quantitative manner and a formula for the intensity of rumor may then be written as follows: $R \sim i \times a$. This formula means that the relation between the two factors is not additive but multiplicative, for if either importance or ambiguity is zero, there is no rumor.

Allport and Postman present this law as "highly dependable," but they inform us that there are certain conditions "under which its operation will be weakened" (p. 34). For instance, if heavy penalties are placed on rumor spreading as, for instance, in Gestapo Germany, or if social barriers prevent crossing. Another reason for the failure of the law to operate may be that a person in the rumor chain knows the rumor law. It is a fact "too little observed by psychologists, that knowledge of the operation of a law frequently alters, and sometimes negates, the law in question" (p. 35).

If this is true, and quite certainly it is, then the formula of the rumor law needs completion. For one should be able to state the weakening conditions of the law in the formula itself. There is another factor which plays a role in the transmission of rumor and it should be involved in the formula. For the weakening conditions are not directly concerned with the factors of ambiguity and importance. These two factors remain constant in a given situation of rumor even though we may know the law. Neither the ambiguity nor the importance of the rumor-fact decreases (e.g., a rumor about the death of Stalin remains at a given moment as ambiguous and as important as it is), even though I know the rumor law.

Therefore in the formula a new factor should be stated to account for the individual disposition of the rumor listener. If this listener knows the rumor law, then that factor will weaken the working of the two other factors, possibly totally negate their working. But, one could object, should a formula be altered for the incidental case in which a person knows that formula? That is an exception and moreover most formulas for psychological laws would have to be altered. Indeed, if it were only for the rare case of a person who knows the psychological law, it might seem silly to alter it. But there is more: the knowledge of the law is only one factor in the individual disposition that may influence the two other factors. It is an expression of something more general, viz., the *critical sense* of the rumor transmitter. In proportion to the critical sense of the individual, the operating of the other two factors will be weakened. The critical sense may be sharpened not only through knowledge of this psychological law but also through psychological insight in general, through perspicacity in social happenings, through higher intelligence, through greater resistance against mass suggestion. Therefore, it seems necessary to change the rumor law as follows: $R \sim i \times a \times \frac{1}{c}$.

Here c stands for critical sense, which contains not only perspicacity but also a volitional factor in the sense of will-power or the disposition to make decisions after reflection and consideration and not to give way immediately to the two mentioned rumor factors. And c also stands for moral criticism or conscience which, in every individual who can bear responsibility, operates as an inhibiting factor in the telling of rumors. If, for instance, I hear the rumor that one of my best friends is an imposter I shall refrain from telling it to others; at any rate I shouldn't tell it to everybody. So c stands for the critical factor in every rumor chain. For rumor is not a wholly mechanical process but a human happening wherein every individual is tested on his critical and moral values.

As c grows greater the rumor will weaken or be totally stopped: the rumor, or the intensity or mobility of the rumor, is inversely proportional to factors i and a. If c equals 1 (i.e., a supposed general average), then the fraction equals 1, which means that the rumor spreads according to the general law without special barriers: $R \sim i \times a \times \frac{1}{1}$. If c is less than average, then the rumor grows stronger. In such a case the rumor transmitter will say, "I have seen it myself" or "my wife saw it happening," to enhance the positiveness of his telling. If c is greater than normal, then the intensity of the rumor will be approaching zero and perhaps it will stop at that link in the rumor chain.

The factor c is, of course, not constant for a given individual in all situations. It, too, varies with the ambiguity and the importance. The critical sense of an anti-Semite, for example, will be weaker if he hears an anti-Semitic rumor than if he hears an anti-English rumor.

The three factors in rumor transmission cover and stand quantitatively in multiplicative relation in the sense of the proposed formula.

REFERENCE

1. ALLPORT, G. W., & POSTMAN, L. *The psychology of rumor.* New York: Holt, 1947.

Received April 17, 1952.

REVIEWS

TOWARD A GENERAL THEORY OF ACTION. By Talcott Parsons and Edward A. Shils (Eds.), Edward C. Tolman, Gordon W. Allport, Clyde Kluckhohn, Henry A. Murray, Robert R. Sears, Richard C. Sheldon, and Samuel A. Stouffer. Cambridge: Harvard University Press, 1951. Pp. xi + 506. $7.50.

Proclamations of the unity of the social sciences are hardly news. At the level of research operations, psychologists, sociologists, and anthropologists undoubtedly share an increasing ground of common interest and understanding; from a more conceptual perspective the *Zeitgeist* clearly has it that personality, society, and culture are too closely enmeshed with one another to be pursued in complacent departmental isolation. How unity emerges from this conceptual trinity, however, remains the subject of conflicting sectarian dogma, and the history of attempts to erect a unified structure by grand legislation is marked by scholastic sterility. When such a distinguished and varied group of authors as those responsible for the present volume undertake to lay out the groundwork of a "general theory of action" into which each of these foci is to be articulated in its proper place, social scientists are justified in viewing the enterprise with a mixture of interest, hope, and misgiving. All of these reactions seem to me to be warranted by this ambitious, difficult, rewarding, and exasperating book.

If the book were in fact the integrated treatise which the reader might expect between covers bearing its title, a review could properly ignore how it came to be written. Such, however, is hardly the case, and one must know the circumstances related in the preface by Parsons to grasp the rationale for its contents. In the fall of 1949, Tolman and Shils were brought to the Department of Social Relations at Harvard to participate in a "stocktaking" of the theoretical resources of the fields that had been joined in the department some three years previously. Parsons and Sheldon of the Department completed the project's full-time staff, which met for a semester in weekly discussions with the other five collaborators in the present volume.

Only those parts of the book written by the four staff members are a direct outcome of the central enterprise. The core of the book, entitled "Values, Motives, and Systems of Action," appears as Part 2 under the authorship of Parsons and Shils. This monograph, originally drafted by Parsons and extensively revised by Shils, contains a detailed exposition of variables and categories articulating personality, culture, and the social system in a comprehensive framework. Part 1, ambitiously headed "The General Theory of Action," contains two chapters. The first, "Some Fundamental Categories of the Theory of Action: A General Statement," bears the joint signature of all nine collaborators in the volume. It is no surprise to read in the preface, however, that it was originally drafted by Parsons as an introductory chapter of Part 2; while it was extensively revised on the basis of group discussion and individual critical comment, the sense in which the resulting document is a group product must probably be qualified in the light of its origin. In the second chapter, "Some Observations on Theory in the Social Sciences," Sheldon contributes what is essentially a trenchant review of the whole enterprise, which is hardly disarming to encounter before one is fully involved in the considerable job of reading the book.

Part 3, "A Psychological Model," stands by itself as a monographic presentation of Tolman's present conceptual system. Tolman shows serious commitment to the common enterprise, nonetheless, in accepting the task of stating in his own terms the equivalents of the key Parsons-Shils variables. Similar commitment is less evident in the papers by the non-staff collaborators that make up Part 4 under the misleading title, "The Theory of Action and its Application." The main reason, so far as I can see, for including these individually interesting contributions in the same volume is that their authors participated in the small discussion group against which the editors whetted their developing conceptualizations. The papers on prejudice by Allport, on social behavior and personality development by Sears, and on role conflict and personality by Stouffer have previously been published in professional journals, and bear no direct relation to the conceptual scheme developed in the general statement and in the Parsons-Shils monograph. The remaining chapters, by Kluckhohn and Murray respectively, are first published here. But neither is an "application" of the earlier conceptual scheme, and each diverges in major respects from it.

The book as a whole, thus, must be said to reflect not the unity of the social sciences but the healthy diversity of the Harvard department. Beyond questioning the judgment of bringing it out as a single work and a group product, and noting a certain pretentiousness in the claims identifying it with *the* theory of action, there is little that a reviewer can appropriately say about the entire volume. The several contributions require consideration in their own right.

In the centerpiece of the book by Parsons and Shils, of which the joint "general statement" is

essentially a selective and simplified version, the authors start with conceiving an "actor" striving to achieve goals. *Action* (which corresponds closely to *behavior* in Tolman's sense, as the latter notes) involves orientation toward objects, in that "it is guided by the meaning which the actor attaches to its relationship to his goals and interests" (p. 4). The "action frame of reference" is thus subjective: concepts are formulated and situations defined from the point of view of the individual actor. Orientation to objects entails selection or discrimination, and this, according to the authors, occurs in three modes: *cognitive,* in regard to location and characterization; *cathectic,* in terms of positive or negative relevance to the satisfaction of drives; and *evaluative,* in so far as selection must be made among alternative objects and gratifications in terms of some explicit or implicit standards. These three modes are grouped together as the *motivational orientation* of the actor; the evaluative standards themselves are said to comprise his *value-orientation.* Value-orientation is then treated as coordinate with motivational orientation, and in turn dissected into three subsidiary modes: *cognitive,* involving standards of truth and validity; *appreciative,* involving standards of the desirable and gratifying that derive from the cathectic motivational orientation; and *moral,* involving "comprehensive integrative standards for assessing and regulating the *entire system of action under consideration,* whether it be a personality or a society or a sub-system of either" (pp. 73–74, italics theirs). If this sounds intricate, so it is, and it is only the beginning.

From these modes of orientation together with a classification of objects of orientation, Parsons and Shils derive a scheme of five so-called *pattern-variables* for the analysis and description of systems of personality, society, and culture. Passages such as the following testify to the great importance which the authors ascribe to this particular formulation:

> . . . There is a certain range of *problems* of orientation which are inherent in the structure of systems of action and . . . an orientation to each of these problems is a functional imperative of action. . . . Since the problems have a determinate form arising from the nature of action, the number and logical relations of the types of alternative solutions is also determinate. Each of the pattern variables states a set of possibilities of selective response to the alternatives presented by the situation of action. We have enumerated five such pattern variables and we have given reasons for believing that it is legitimate to consider them an exhaustive set. The exhaustive character of the classification of pattern variables has far-reaching implications for the analysis of systems of moral standards . . . (p. 171).

> A *pattern variable* is a dichotomy one side of which must be chosen by an actor before the meaning of a situation is determinate for him, and thus before he can act with respect to that situation. We maintain that there are only five *basic* pattern variables . . . and that, in the sense that they are *all* of the pattern variables which so derive, they constitute a system (p. 77, italics theirs).

The five dichotomous pattern variables can at least be identified here. They are:

1. Affectivity—Affective neutrality (depending on whether or not cognitive-cathectic modes determine behavior prior to evaluation).
2. Universalism—Particularism (depending on whether cognitive or appreciative standards have primacy).
3. Self-orientation—Collectivity-orientation (depending on the absence or presence of evaluation in terms of collective moral standards).
4. Ascription—Achievement (depending on whether the actor perceives a social object in terms of qualities or performances).
5. Specificity—Diffuseness (depending on whether or not the rights of a social object in relation to the subject are limited by explicit definition).

In cross classification they generate 32 cells. This highly abstract schematization provides the basic framework for classifying need-dispositions of personality and role-expectations in social systems with respect to their value components, that is, with respect to the standards of selectivity that operate in the actor's orientation. In the description of personality, diffuseness-specificity and neutrality-affectivity are given focal position, while universalism-particularism and ascription-achievement play the corresponding primary role in the analysis of social systems. The entire gamut of possibilities comes into play in the treatment of cultural systems of value orientation.

The pattern variables and their permutations do not, of course, exhaust what Parsons and Shils have to say about systems of society, personality, and culture. But they clearly consider this scheme to be their principal contribution. Since it purports to be based on an examination of the elements of behavior, social psychologists may well put it to some searching questions. For a start, here are mine:

1. *Are the pattern variables in fact an exhaustive list?* As an analytical formulation, the set of variables may well embrace the logical possibilities of priority among the modes of orientation and types of social objects which Parsons and Shils have chosen to recognize in their analysis of action, though it is possible that other theorists taking the same starting point might employ different principles of grouping and arrive at distinguishably different results. The real question hinges on the analysis of action from which they are derived. The list is just as fictional, just as arbitrary, as these assumed elements of orientation and situation. To me, a high degree of arbitrariness seems to be involved in selecting the quality vs. performance and specificity vs. diffuseness distinctions as the sole and fundamental ways of categorizing perceived social objects. Yet any of the many other possible ways of classifying them that seem equally compatible with the rest of the "action frame of reference" would generate a different list of pattern variables.

2. *Are the pattern variables appropriately regarded as dichotomous?* Their cross classification to generate neat descriptive categories requires them to be so. Yet the assumption that they are not continua seems to me gratuitous, one that squares poorly with the working assumptions of psychologists and, for that matter, is ill adapted to existing techniques of empirical investigation.

3. *Do the systems of classification which they generate "fit" the needs of personality, social, and cultural theory?*

A classification is not true or false; apart from matters of consistency it is to be judged in terms of whether it highlights distinctions that are important on other grounds. In the cross classification of need-dispositions according to the pattern variables (p. 250), one notes, to pick instance at random, that affective neutrality combined with particularism, diffuseness, and ascription yields "disposition to esteem and want to be esteemed by a particular object possessing specific qualities." It is not made clear why psychologists or others should concern themselves with such distinctions. The relevance of the pattern variables to Parson's analysis of social systems should be greater, as he had already made considerable theoretical use of them before arriving at their present derivation.

4. *Do the pattern variables contribute to a satisfactory general theory of social behavior?* In themselves, they do not constitute a theory, since they make no assertion about empirical fact. Unfortunately, the task Parsons and Shils have set themselves in their chapters does not include the framing of testable propostions in terms of these or other variables; one misses the stress on antecedent-consequent relations that characterizes Sears' paper. Neither is attention given to the problem of how the variables are to be given operational meaning. One can therefore only wait and see whether they turn out to be useful in the program of theoretically oriented research which the reader is given to understand is under way. This point is made by Sheldon in Chapter 2.

The attempt to construct descriptive categories with highly general applicability is obviously valuable. The danger of the present formulation lies in its pretension to finality and completeness, and the degree to which it carries conceptual elaboration in the absence of data.

While the pattern variables are a major unifying theme in the chapters of Part 2, other contributions seem to me to have more value for social psychology. The analysis of the conceptual status of personality, culture, and the social system in relation to one another, for example, is elegant and persuasive. Both personalities and social systems, say Parsons and Shils in agreement with their collaborators, are concrete systems of action, the first centering on the orientation of any one actor, and the second on the structure of relationships emerging in the interaction of a plurality of actors. The two are at the same conceptual level; while their connection is intimate, "the social system is not a plurality of personalities" (p. 7). Culture patterns, on the other hand, have different status: culture, though "systematic," is not an action system coordinate with the other two types. The authors speak of systems of ideas or beliefs, expressive symbols, and value orientation, "system" in this context referring to the meaningful coherence among cultural elements abstracted from concrete action systems. The cultural tradition enters into action systems either as an object of orientation or as an intrinsic aspect of orientation, the latter in the important cases of internalization in personality and institutionalization in the social system.

If this sorting out of concepts fails to find favor, its lack of correspondence with a job analysis of the activities of contemporary sociologists, personality psychologists, and cultural anthropologists will probably be to blame. Yet there is no reason to suppose that the distinctions essential to an encompassing framework should turn out to correspond to academic boundaries that have grown up by happenstance. "Concrete systems of action," as the General Statement puts it (p. 7), "have psychological, social, and cultural aspects." Distinguishable analytically, they are empirically inseparable.

Two other points, among the many that could be mentioned, struck me as particularly relevant to social psychology. The treatment of the complementary nature of sanctions and role expectations (pp. 15, 154) is incisive, and deserves the close attention of those who regard role as a central concept of social psychology and a link between psychological and sociological theory. The chapter on systems of value orientation, and, indeed, the emphasis on problems of evaluation throughout Part 2, raise important problems which psychologists have been prone to neglect. It should also be added that the treatment of personality shows sophisticated familiarity with psychoanalytically inspired dynamic psychology. The "structural-functional" approach with which Parsons is identified as a sociologist has, of course, strong affinity with the familiar psychological orientation around problems of integration, adaptation, and adjustment.

Turning now to Part 3, we can treat Tolman's psychological model more expeditiously, since earlier versions of his views are familiar to psychologists. Tolman's happy style comes as no small refreshment to the eye of the reader who has traversed the preceding abstract and difficult pages. His model consists in a classification of the independent and intervening variables that he regards as necessary for the prediction of behavior, his dependent variable, together with an indication of their presumed interrelations. His variables contrast, therefore, with those of Parsons and Shils in that their postulated causal relevance is clear from the beginning. There are three classes of independent variables: the stimulus situation, conditions of drive arousal or satiation, and heredity, age, sex, drugs, endocrines, etc., the last of which Tolman dismisses for his present purposes by relating it to capacity and temperamental traits as intervening variables. The other and principal intervening variables, in approximate causal sequence, are: need system, belief-value matrix, immediate behavior space, locomotion in behavior space, and restructured behavior space as a result of locomotions and learning mechanisms. He devotes ten pages to the problem of operationally defining the main attributes of these variables. The total schema can be grossly described as a quasi-Lewinian psychology anchored in principle to causal antecedents and behavioral consequents and freed of Lewin's topological baggage, though

Tolman's unique conceptual diagrams, the contribution of which to clarity of presentation is debatable, could be said to replace the latter. The model serves to order relationships already obtained in behavior research and to suggest directions for further inquiry; it is not presented as a tight or developed theory from which propositions can be derived or predictions drawn.

A model of just this degree of explicitness seems to me to be highly appropriate to the present state of psychological knowledge. I nevertheless found myself slightly disappointed in the results of its application to the psychodynamic mechanisms of identification, the self-ideal, repression, and symbolic substitution, and to perceptual accentuation. Avowedly a venture in *post hoc* interpretation, Tolman's treatment of these problems should not be criticized for being such. But the lack of more highly developed postulates regarding the causal connections assumed in his system left me with the feeling that little has been added to what is already known or hypothesized about these examples. They do serve, however, what was perhaps the only purpose intended: they demonstrate the relevance of Tolman's classification of intervening variables to complex problems of personality and social psychology. It is no small contribution to legitimatize for the behavior theorist, as Tolman does in his final chapter, problems of value standards, social roles, and personality by translating them into the terms of his general system. In his treatment of all of these, the concept of belief-value matrices plays a crucial part.

Of the assorted papers in Part 4, there is space for only the briefest comment on the two that appear for the first time in this volume. Kluckhohn's chapter is a sophisticated attempt to introduce clarity into the current semantically confused discussion of values. The definition on which he elaborates is: "A value is a conception, explicit or implicit, distinctive of an individual or characteristic of a group, of the desirable which influences the selection from available modes, means, and ends of action" (p. 395).

Like the somewhat different treatment of values earlier in the volume, this chapter should prove informative and stimulating to psychologists becoming interested in an important area which they have hitherto neglected. While it puts the reader in touch with a number of currents of thought relevant to value problems, it does not, I think, carry the systematization of the area very far for psychological purposes.

Murray's chapter is disappointing in that his title, "Toward a Classification of Interactions," gives promise of developments that do not occur. After an initial statement that the social scientist's "real entity" is a unit of interaction and that in representing such a unit "the object must be given the same conceptual status as the subject" (p. 435), the chapter turns out to be an up-to-date and forthright statement of Murray's familiar system, in which the primary conceptual problem is the classification of needs or tendencies. Interaction seems somehow to drop out of the picture. Murray does not seem to have been as deeply infected with the social relations virus as he would have the reader believe: how else can one account for his simply equating the concepts of role and function (the latter in the sense that respiration is a physiological function) in a section entitled "role behaviors"?

Once these reservations are set aside, the chapter is a rewarding one, an eloquent and persuasive plea for a dynamic, humanistic psychology unencumbered by rat-born models. Some developments in Murray's thinking may be noted in passing. Introducing the neologism *proaction* as contrasted with reaction, he suggests that such emitted behavior is not merely homeostatic and restorative of equilibrium. Murray's formulations offer no convenient way out of the tension-reduction formula, but they place him among those who are giving genuine problematic status to the spontaneous, constructive, creative aspects of behavior. At the end of the chapter, an outline is presented of a revision of his list of needs in terms of new lists of vectors and values, taken in cross classification with one another. The exposition is too brief to give much indication of the descriptive or theoretical utility of the scheme, which again leaves me with misgivings about the fruits of preoccupation with systems of classification and cross classification. But the restructuring brought about in the original list of needs by this bi-dimensional approach would seem to be promising.

The book, in sum, is one which many psychologists will want to consult for its component parts. To read it from cover to cover, as few are likely to do at one set of sittings, is an arduous but educational experience. It is, moreover, an experience that should contribute genuinely toward that integration of the social sciences which is its aim. Not, to be sure, that it offers a single final framework which one has only to adopt; rather, it exposes the reader to serious thinking from a variety of perspectives on vital problems of interdisciplinary concern.

M. BREWSTER SMITH

New York City

AN INTRODUCTION TO PROJECTIVE TECHNIQUES AND OTHER DEVICES FOR UNDERSTANDING THE DYNAMICS OF HUMAN BEHAVIOR. Edited by Harold H. Anderson and Gladys L. Anderson. New York: Prentice-Hall, 1951. Pp. xxiv+720. $6.75.

According to the editors the goal of this book is "to afford a general survey of the field of projective

techniques. . . ." In the opening chapter, there is recognition by one of the editors (H.H.A.) that the term *projection* has come to mean many things to many people, but he sees the classical definition a la Freud as a point of orientation for "most of the contributors to this book." That the editing of the book has erred on the side of liberality is suggested by the observation of H. A. Murray in the foreword: "Of projection, according to Freud, I have seen in this collection of papers, practically no evidence."

It would likely have proved unhappily cramping to the stated purpose of the book if the editors had selected for inclusion only those techniques whose exponents appeared to subscribe to a single concept of projection. However, in view of the several meanings of projection as applied to instruments for personality study and of the heterogeneity of the meanings and techniques subscribed to by the writers appearing in this volume, there is need for a chapter devoted to delineation of the various facets of projection and concerned with a logical ordering of the multitude of procedures which have come to be labeled (or mislabeled) as "projective" techniques. There is such a chapter, by R. B. Cattell, but one may be disturbed to find that it is the third chapter of the book, that it follows rather than precedes an excellent discussion of validation problems (Chapter II by Jean Walker Macfarlane and Read D. Tuddenham) and that the exposition of "Design Principles in Projective Techniques" seems somewhat subservient to an account of the construction, reliability, and validity of some of Cattell's paper-and-pencil "dynaception tests." Nevertheless, Cattell should be thanked for providing the only systematic development of the psychology of projection in relation to the design of techniques in general, and the editors must be complimented for courage to fly in the face of prevailing cultural winds by including in a volume on projectives a chapter that proposes to measure dynamisms with what superficially appear to be those most unpopular of devices, so-called "structured" tests.

What one would ordinarily expect to be the introductory chapter, Chapter One, presents only a fleeting reference to the definitional problem and then develops a theoretical discussion of personality growth. This discussion is pertinent, of course, to projective techniques, as any account of personality development would be. However, the discussion is a general one and not clearly tied to the problems presented by projective theory and devices. In this sense it seems almost detached from the major content of the book and does not properly or adequately serve an introductory function. The lack of an orientational chapter concerned with the definition of terms, a statement of fundamental psychological principles, and a general review of current status might well have been offset by a summary chapter but such a summary is also lacking.

More positively, it may be said that the three chapters which constitute Part I, those by H. H. Anderson, Macfarlane and Tuddenham, and Cattell, constitute a more thorough and more critical, albeit somewhat disjointed, introduction to the subject of projective techniques than has been provided in other surveys.

There are four other parts to the volume. Part II, devoted to the Rorschach, includes an expository chapter by Beck and a review of experimental studies by Rabin. Part III includes chapters on other techniques including the Four-Picture Test, TAT, Word Association and Sentence Completion, Rosenzweig P-F Study, Bender-Gestalt, Draw-a-Human, Finger Painting, Graphology, Expressive Movement, and the Szondi. A single chapter reviews lesser known techniques such as Sargent's Paper-Pencil Projective Method, story telling and story completion methods, the H-T-P technique, the World Test, and the Tautophone Method.

Part IV, General Intelligence Tests in Personality Appraisal, includes chapters on the Wechsler-Bellevue and Stanford-Binet, the latter by one of the co-editors (G.L.A.). Finally, Part V is concerned with projective techniques as media for therapy and includes chapters on puppetry, play, and psychodrama.

In a work involving, as this one does, a great range of materials and a number of writers, it is reasonable to expect considerable variability in the level of writing, the objectivity of attitude, and the thoroughness of coverage. Something of the unevenness of scholarship is attested by the fact that of the three writers who trace the origin of "projection" to Freud, the sources cited in each instance are different and bear the dates 1894, 1933, and 1938 respectively! It is interesting that R. R. Holt in developing the rationale of the TAT does not concern himself with the origins of projection as a recognized dynamism but does draw parallels between Freud's analysis of the "alchemy" of dreams and the process of theme formation.

Twenty-three of the twenty-four chapters have reference lists the average length of which is 30 items. In view of the mountain of publications on projectives, this volume obviously was not intended, or at least will not serve adequately, as a reference work. For example, the chapter by Rabin, which is supposed to review validating and experimental studies with the Rorschach, cites only 37 references. Holt's chapter on the Thematic Apperception Test carries but 43 references. Clearly there has been selection, and from the general tenor of these and similar chapters clearly the selection has been pro-projective. This is not entirely inappropriate in a book of this nature and one assumes that the limited citations given by the various authors are at least partly reflective of editorial limitations on space. Inasmuch as scholarly, exhaustive examination of all pertinent literature was prohibited by general space limitations and similar considerations, the

editors might have sought a degree of impartial reporting by selecting as their writers on the various devices individuals having no prominent identification with them and hence no likely proclivity for axe-grinding. This was not done, and most of the techniques are reported by workers having marked investments, e.g., Beck on the Rorschach, van Lennep on the Four-Picture Test, Machover on the Draw-a-Human technique, Mayman, Rapaport, and Schafer on the Wechsler, Wolff on graphology, and Moreno on psychodrama. Again, this is not necessarily inappropriate or undesirable, but the professional reader or student should be aware of these selectivities so as not to assume that he is being served the "whole truth and nothing but the truth."

On the other hand, this reviewer did not detect the general aura of extravagance which has characterized so much of the polemics of projectivists. There is very little evidence from any of the authors of an attitude of *"This* is the technique nonpareil." One wonders if this sweet reasonableness results primarily from editorial policy which discouraged exhortation, or if it reflects the current attitudes of projectivists who may recognize the precarious status of projective devices which are most secure in terms of universality of popular use and most insecure in terms of convincing evidence of usefulness. In this regard, the general rationale argued for the "tests of expression" presented in this book gives them all of the inherent validity and all of the actual predictive usefulness of an assertion, "All men will die." With respect to the possible influence of editorial reserve in setting the tone of the book, we know only that with but two exceptions (Chapter 1 by H. H. Anderson on personality growth and Chapter 24 by D. M. Levy on hostility patterns) each chapter was "prepared specifically for this book."

The goal of this book is to survey the field of projective techniques. It is, by title, an *introduction.* The goal has been well achieved, and the title is accurate. The editors are to be complimented for their selection of the techniques which are represented, a selection process made difficult by the amoeba-like birth rate of projectives. The book will serve well for introductory survey courses. It is of limited value for practicum courses, inadequate as a text for advanced graduate students, and does not provide a comprehensive reference work. It does a good job of what its editors intended it to do. It is hardly a criticism of the editors that they have not given us what is most sorely needed in this field at this time, namely, a thorough critique of the psychology of projection and an exhaustive analysis of the current *scientific* status of projective devices.

WILLIAM SCHOFIELD
University of Minnesota

ABNORMAL PSYCHOLOGY. By David B. Klein. New York: Holt, 1951. Pp. xvii+589. $4.75.

Klein's response to the urgings of E. S. Conklin's friends and publishers who are eager to serve the present generation of college students a generous, up-to-date helping of the latter's *Principles of Abnormal Psychology* (1935), has resulted in an inferior rehash of the usual fare currently being offered to students of abnormal psychology. The present product is more palatable than the earlier edition, but it lacks a great many of the original's basic ingredients. It adds, instead, many morsels of contemporary thinking which have been prepared more impressively by other textbook writers also interested in the student's diet of abnormal psychology.

Except for two or three chapters, this book has been written independently of the earlier text, and represents a "radical shift in organization, emphasis and even in viewpoint with respect to some topics" (p. vi). While Conklin thought of abnormal psychology as a "basic science" (in both his 1927 and 1935 volumes), Klein orients his treatment of the subject chiefly in the direction of those students who are planning to apply their knowledge in the clinical area. In the eclectic spirit of Conklin, he maintains in this text that the exposition of abnormal psychology "in terms of a single systematic viewpoint . . . could be done only at the cost of a dogmatic disregard of much that fails to square with the school selected" (p. 22). Nevertheless, he does rely heavily on the psychodynamic interpretation of abnormal behavior as contributed by Freudian and neo-Freudian viewpoints, thereby going beyond Conklin's guarded acknowledgment of the possible value of these viewpoints. The focal points in Klein's presentation are the "insecurity syndrome" and the adjustive maneuvers of the individual aroused by the need to keep self-esteem intact.

> As has been repeatedly brought out on previous pages, to a very large extent so-called "functional" disorders are products of conflicts involving the insecurity syndrome. Over and over again such conflicts were seen to be derivatives of the individual's quest for affection and admiration, and all that comes to influence the self-regarding sentiment. In this sense the functional disturbances were seen to involve interpersonal difficulties, the understanding of which called for more of a social science orientation than the physico-chemical approach of traditional medicine (p. 425).

By and large Klein's exposition is lucid enough, and his examples and illustrations are frequently apt, imaginative, and chosen with his reader in mind. He makes abundant use of section and subsection headings, and demonstrates a flair for selecting headings which sometimes border on the dramatic. In the nineteen chapters of the book, he seems to touch on practically all the material that

might conceivably fall within the realm of abnormal psychology, including three chapters (104 pages) on therapy. There are 28 figures, over half of which will doubtless appeal to and capture the interest of the average student. Finally, he includes "A Review Outline of Psychopathology" (Chapter 16) which, as he recommends, may be used both as an introduction to and a summary of most of the material in the text.

In spite of these features, however, the book leaves much to be desired both as an exposition of the subject matter of abnormal psychology and as an attempt to furnish the budding clinical psychologist with a conceptual framework in which to develop, a chief goal which Klein has intended to pursue. Not infrequently important points are brought up which are left dangling, flimsily developed, or otherwise dissociated from the general direction that the discourse seems to be taking. Not infrequently side issues, apologies, and cautions are interspersed in the development of themes, tending to leave the reader more with an awareness of the complexities and confusions in the subject matter than with a feeling that most of the more or less salient points were presented. Not infrequently, in exercising the understandable prerogatives of authorship, Klein plays up or plays down topics in directions which would appear to be contrary to the needs of the average student, especially if he expects to major in the clinical area. These impressions might be clarified by brief references to several specific instances.

Although he insists that "the hallmark of a neurosis is the existence of anxiety" (p. 434) and "the ramifications of anxiety can become so widespread as to make for mental illness" (p. 48), there is nowhere to be found a discussion of the meaning of anxiety, its behavioral characteristics, the conditions under which it develops, or the manner in which it results in disordered behavior. Similarly, the "Somatization Reactions" or psychosomatic disorders are catalogued in about two pages (which, incidentally, are not in accord with the page reference given in the index), and are described in such a superficial and fragmentary manner as to create a Zeigarnik effect, to borrow a feeling from Klein, which the interested student will have to dispel by referring to other sources. Chapter 3, which deals with "the self and its dynamisms" and which would appear to be important to his psychodynamic orientation, is rather disjointed and interspersed with forced, parenthetical issues. In ordering the material of this chapter, his section headings read, among others, "dynamisms of blame-avoidance," "dynamisms of blame-acceptance," and *"some defense dynamisms"* (italics mine), but nowhere is there any definitive discussion of repression as a dynamism! Chapter 17, which presumably deals with the current variations in psychoanalytic psychotherapy, seems to be more a descriptive summary of the history and current controversies in psychoanalytic theory than a successful effort to "abstract the seemingly tenable features of the rival schools of psychotherapy" (p. viii).

In his discussion of clinical tests of impaired thinking, Klein dwells considerably on what he has christened the "Columbia test," and he suggests in a later chapter that it be employed, in conjunction with the Vigotsky and various sorting tests, to reveal "latent schizophrenia" (pp. 452–453). Actually this "test" was an *ad hoc* technique used by Rashkis, Cushman, and Landis in a study of the thinking process of only 36 subjects, including 8 schizophrenics and 10 paretics, and it does not enjoy a wide use among clinical psychologists. Moreover, ascertaining the diagnosis of latent schizophrenia by the use of such instruments is still considered risky. Hence, it would appear that the student might arrive at a distorted perspective of the types of tests used in clinical practice and the value of such tests in deciding diagnosis.

In summary, it may be said of Klein, in Kipling's words, "what 'e thought 'e might require 'e went an' took." The product of his endeavors, however, seems to be divested of originality in viewpoint, impressiveness in presentation, and inspiration in organization and integration of the compatible elements of recent evidence and theorizing. In this reviewer's opinion, therefore, this book is not likely to replace or even challenge those books which currently enjoy important places in the classes of abnormal psychology.

BERNARD SAPER

Austin State Hospital

AIR WAR AND EMOTIONAL STRESS. By Irving L. Janis. McGraw-Hill, 1951. Pp. ix+280. $5.00.

As a research consultant to the RAND Corporation, under whose sponsorship this volume appears, Janis prepared a series of memoranda on the psychological effects of air war. On the basis of these reports, he wrote the present book which explores the implications of available findings for over-all military strategy, for civil defense, and for the political, social, and moral consequences of atomic warfare. After surveying the implications of the data which have already been collected, Janis spells out the questions which these data leave unanswered and suggests, in considerable detail, a program of research which will supply us with the information needed to establish a sound civil defense policy.

Janis begins with an evaluation of the emotional impact of the A-bomb and the aftereffects of atomic bombing, basing this study on the United States Strategic Bombing Survey of Japan, on John Hershey's *Hiroshima* and on various eyewitness accounts of the Hiroshima and Nagasaki bombings. From this he proceeds to a more general study of

the effects of air war, focusing mainly on the effects of bombing in England and Germany. Again, he relies to a great extent on the USSBS German report but, in this instance, he has at his disposal a considerably greater amount of supporting data in the form of reports of social scientists, psychiatrists, administrators, and reporters. Janis investigates the effect of air war on psychiatric disorders, on aggression and morale, and the patterns of adaptive behavior whereby the individual accommodates to air war.

In his concluding section on civil defense, he discusses topics ranging from the types of bomb shelters that should—or *can*—be constructed, the motivation of civil defense workers, optimum civil defense policy from the standpoint of civil morale in wartime, psychological selection of civil defense workers, education of the public for civil defense, including the proper timing of such an educational program.

Obviously, the range and number of discrete problems covered in this book are so great that the conclusions and policy suggestions at which Janis arrives cannot be summarized within the scope of a review. This much can be said however—and this is what makes summarization so impossible—Janis' conclusions, both substantive and policy, are characterized by a concreteness of thinking that undercuts and often reverses many generalizations that have been accepted on the basis of fewer data or less rigorous analysis. On the policy side, for example, he suggests the dual dangers of inaugurating a civil defense policy when danger is too remote (in which case the public will fail to be aroused), or in a period of crisis (in which case there will be either a reaction of apathy—"it is too late to do anything"—or an emotional letdown with a feeling that the government has been crying "wolf" if the crisis passes over). ". . . the most favorable time for introducing a full-scale civil defense program and for demanding all-out public participation would be during a period when war does *not* appear to be imminent but during which public attention is focused more and more upon the threat looming up on the distant horizon."

As an example of applied social psychology directed to problems of national importance done during peacetime, this book is important for a variety of reasons. It is my impression (despite Dr. Janis' expressed hope to the contrary) that very little of importance to psychology as a discipline comes out of this book. The contribution is almost exclusively in the other direction. The only contributions which Janis has to make to the science of psychology are in pointing out those gaps in our systematic knowledge which make it impossible to answer certain questions of a practical nature. This is primarily an applied work, and as such it demonstrates that the psychologist has a good deal to contribute to practical problems of national policy, both by virtue of the theoretical and empirical propositions his science has accumulated, and as a result of his training in pursuing ruthlessly the implications of his data.

Perhaps the most important single policy suggestion that comes from Janis' work is one which he never makes, but which the reader cannot escape. It stems from the great danger of doing social research under conditions which force one to skim off the cream of his conclusions without appropriately digesting his data. The USSBS reports were excellent jobs considering the circumstances under which they were written, and the surveys on which they are based furnish a mountain of valuable data. Some of the conclusions of these reports have been attacked previously and Janis, though he relies heavily on these studies, constantly pushes beyond and challenges the USSBS conclusions, often on the basis of data in the reports themselves. It is not to the derogation of the writers of those reports to say that under less pressure of time they might have come to different conclusions; nor is it to be expected that administrators will not, in the future, insist on hurry-up results in a time of crisis. These circumstances do suggest, however, the need for continuing peacetime research of the high caliber exhibited in this book on problems of national importance. If this is done, the forced-draft work of a crisis period can be related to, and checked against a body of better digested, more thoroughly thought-through findings.

RAYMOND A. BAUER
Harvard University.

NEW BOOKS RECEIVED

ABRAMSON, HAROLD A. (Ed.) *Problems of consciousness.* New York: Josiah Macy, Jr., 1952. Pp. 156. Price $3.25.

BAL, ALEXANDRE. *L'Attention et ses maladies.* Paris: Presses universitaires de France, 1952. Pp. 123. Price not supplied by publisher.

BARNETT, GORDON J., HANDELSMAN, I., *et al. The occupational level scale as a measure of drive.* Psychological Monographs, Vol. 65, No. 10 (Whole No. 342). Washington: Amer. Psychol. Ass., 1952. Pp. 37. Price $1.50.

BAUMGARTEN, FRANZISKA. *Zu spat . . . von der tragik menschlichen handelns.* Zurich: Rascher, 1952. Pp. ix+150. Price not supplied.

BELLAK, LEOPOLD. *Psychology of physical illness.* New York: Grune & Stratton, 1952. Pp. vii+243. Price $5.50.

BELLAK, LEOPOLD. *The supplement to the children's apperception test.* New York: C.P.S. Co., P.O. Box 42, 1952. No pagination. Price not supplied.

BENNETT, M. E. *College and life: problems of self-discovery and self-direction.* New York: McGraw-Hill, 1952. Pp. ix+457. Price $4.50.

BETT, W. R. *The infirmities of genius.* New York: Philosophical Library, 1952. Pp. 192. Price $4.75.

BISNO, HERBERT. *The philosophy of social work.* Washington: Public Affairs Press, 1952. Pp. x+143. Price $3.25.

DEESE, JAMES. *The psychology of learning.* New York: McGraw Hill, 1952. Pp. vii+398. Price $5.50.

EISSLER, RUTH S., FREUD, ANNA, *et al. The psychoanalytic study of the child—VII.* New York: International Universities Press, 1952. Pp. 448. Price $7.50.

ELLSON, HAL. *The golden spike.* New York: Ballantine Books, 1952. Pp. 243. Price $.35 (paper-bound).

FIEDLER, MIRIAM FORSTER. *Deaf children in a hearing world.* New York: Ronald, 1952. Pp. viii+320. Price $5.00.

FRENCH, DAVID G. (Ed.) *The social welfare forum, 1952.* New York: Columbia Univer., 1952. Pp. xviii+305. Price $4.75.

FREUD, SIGMUND. *An autobiographical study.* New York: Norton, 1952. Pp. vi+141. Price $2.50.

GILBERT, JEANNE G. *Understanding old age.* New York: Ronald, 1952. Pp. ix+422. Price $5.00.

GORLOW, LEON, HOCH, ERASMUS L., & TELSCHOW, EARL F. *The nature of non-directive group psychotherapy.* New York: Teachers College, Columbia University, 1952. Price $3.25.

GUTTMACHER, MANFRED S., & WEIHOFEN, HENRY. *Psychiatry and the law.* New York: Norton, 1952. Pp. viii+476. Price $7.50.

HALL, BERNARD H., GANGEMI, MARY, *et al. Psychiatric aide education.* New York: Grune & Stratton, 1952. Pp. xvi+168. Price $5.75.

HOYLES, J. ARTHUR. *The treatment of the young delinquent.* New York: Philosophical Library, 1952. Pp. viii+273. Price $4.75.

HULL, CLARK L. *A behavior system—an introduction to behavior theory concerning the individual organism.* New Haven: Yale Univer. Press, 1952. Pp. ix+372. Price $6.00.

JOHNS, ALFRED E. *Scientific autosuggestion for personality adjustment and development.* New York: Associated Booksellers, 1952. Pp. 104. Price $2.95.

JOHNSON, WENDELL, DARLEY, FREDERIC L., & SPRIESTERSBACH, D. C. *Diagnostic manual in speech correction.* New York: Harper, 1952. Pp. viii+221. Price $2.50.

KECSKEMETI, PAUL. *Meaning, communication, and value.* Chicago: Univer. of Chicago Press, 1952. Pp. viii+349. Price $8.50.

KOHLER, FRED. *Evolution and human destiny.* New York: Philosophical Library, 1952. Pp. 120. Price $2.75.

KUHLEN, RAYMOND G., & THOMPSON, GEORGE G. (Eds.) *Psychological studies of human development.* New York: Appleton-Century, 1952. Pp. xii+533. Price $3.50.

LANDIS, PAUL H. *Adolescence and youth.* New York: McGraw-Hill, 1952. Pp. xii+461. Price $5.00.

LE CRON, LESLIE M. *Experimental hypnosis.* New York: Macmillan, 1952. Pp. xiv+483. Price $6.00.

MCKILLIP, ANNE SELLEY. *The relationship between the reader's attitude and certain types of reading response.* New York: Teachers College, Columbia University, 1952. Pp. viii+101. Price $2.75.

MERTON, ROBERT K., FISKE, MARJORIE, & KENDALL, PATRICIA. *The focussed interview.* New York: Bureau of Applied Social Research, Columbia Univer., 1952. Pp. xxv+202. Price $3.00.

MEYERS, ALAN S. (Ed.) *Social and psychological factors in opiate addiction.* New York: Bureau of Applied Social Research, Columbia Univer., 1952. Pp. 170. Price $1.00.

NEGRI, VITALI. *Your complexes and you.* Los Angeles: Amer. Book Institute, 1952. Pp. x+299. Price $3.50.

POWYS, JOHN COWPER. *The inmates.* New York:

Philosophical Library, 1952. Pp. xi+318. Price $4.50.

PUMPIAN-MINDLIN, E. *Psychoanalysis as science.* Stanford: Stanford Univer. Press, 1952. Pp. x+174. Price $4.25.

REIK, THEODOR. *The secret self.* New York: Farrar, Straus & Young, 1952. Pp. 329. Price $3.50.

ROBACK, A. A. *History of American psychology.* New York: Library Publishers, 1952. Pp. xiv+426. Price $6.00.

SHAFFER, G. WILSON, & LAZARUS, RICHARD S. *Fundamental concepts in clinical psychology.* New York: McGraw-Hill, 1952. Pp. xi+540. Price $6.00.

SJOGREN, TORSTEN. *Genetic-statistical and psychiatric investigations of a west Swedish population.* Copenhagen: Ejner Munksgaard, 1948. Pp. 102. Price not supplied.

SJOGREN, TORSTEN, SJOGREN, HAKON, & LINDGREN, AKE GITT. *Morbus alzheimer and morbus pick—a genetic, clinical and patho-anatomical study.* Copenhagen: Ejnar Munksgaard, 1952. Pp. 152. Price not supplied.

SLAVSON, S. R. *Child psychotherapy.* New York: Columbia Univer. Press, 1952. Pp. xiii+331. Price $4.50.

TRENAMAN, JOSEPH. *Out of step.* New York: Philosophical Library, 1952. Pp. viii+223. Price $4.75.

NEUROLOGY & PSYCHIATRY

Section VIII of Excerpta Medica

The Neurology Section includes abstracts of all articles on neurological surgery in addition to general clinical neuro-physiology.

The section on Psychiatry covers the topics usually contained within the meaning of the word. In addition, the following special branches are dealt with fully:

- ★ **Psychoanalysis, individual and analytical psychology**
- ★ **Sexology, criminology, alcoholism, and drug addiction as related to psychiatry**
- ★ **Social and industrial psychology and psychiatry, vocational guidance and mental testing**
- ★ **Psychometrics**
- ★ **Personality testing and the Rorschach test**
- ★ **Heredity and statistical studies as they apply to psychiatry**
- ★ **Mental defect and epilepsy**

Annual volume of 1000 pages, including an index classified by author and subject.

Subscription price $22.50 per annual volume

Excerpta Medica Service Corporation
280 Madison Ave.
New York 16, N. Y.

Please enter my subscription to Section VIII, Neurology and Psychiatry, of Excerpta Medica. I enclose $22.50.

Name .....

Address.....

.....

Pastoral Psychology

A Religiously Oriented Monthly Periodical of the Science of Human Behavior

SEX AND THE CHURCH

. . . a profound, penetrating exploration of the subject from the historical, theological, psychological, philosophical, psychiatric, and legal aspect—by the most outstanding authorities in this country.

Section I (September 1952)

Sex—Sin or Salvation?
by SEWARD HILTNER
Associate Professor of Pastoral Theology
Federated Theological Faculty
The University of Chicago

Christianity and Sex—An Historical Survey
by ROLAND BAINTON
Professor of Ecclesiastical Church History
Yale University Divinity School

A Pastoral Theology of Sex and Marriage
by REUEL L. HOWE
Professor of Pastoral Theology
Virginia Theological Seminary

Pastoral Problems of Sex
by CARROLL A. WISE
Professor of Pastoral Psychology and Counseling
Garrett Biblical Institute

Psychosomatic Aspects of Love
by DR. GOTTHARD BOOTH, *Psychiatrist*
Commission on Religion and Health
National Council of the Churches of Christ

Sex as an Expression of Personal and Social Value
by LAWRENCE K. FRANK
Director (until recently) of the Caroline Zachry Foundation

Section II (February, 1953)

A Philosophy of Sex
by JOSEPH FLETCHER
Professor of Christian Social Ethics
Episcopal Theological School
Cambridge, Massachusetts

A Psychoanalyst's Viewpoint on Sexual Problems
by DR. JOHN A. P. MILLET
Chief Psychiatrist
Rehabilitation Center
The American Rehabilitation Committee

Toward a Christian View of Sex Education
by PETER A. BERTOCCI
Professor of Philosophy
Boston University

Sex and the Law
by JUDGE MORRIS PLOSCOWE
Executive Director
Committee on Organized Crime
American Bar Association

Don't Miss These All-Important Issues!

PASTORAL PSYCHOLOGY
Great Neck, N. Y.

I enclose $3.00 for a one-year subscription
$6.00 for a three-year subscription
(Canadian and foreign subscriptions 50 cents per year additional)

Name .....

Address .....

City ..... Zone..... State.....

Journal of Abnormal and Social Psychology

AVAILABLE BACK ISSUES

Year	Volume	Available Numbers	Price per Number	Price per Volume
1906–07	1	— — — — — —	----	----
1907–08	2	— — 3 — — 6	$1.50	$3.00
1908–09	3	— — — — — —	----	----
1909–10	4	— — — — — —	----	----
1910–11	5	— 2 3 — — 6	$1.50	$4.50
1911–12	6	— — — — — —	----	----
1912–13	7	— — — — — —	----	----
1913–14	8	— — — — — —	----	----
1914–15	9	— — — — — —	----	----
1915–16	10	— — — — — —	----	----
1916–17	11	— — — — — —	----	----
1917–18	12	— — — — — —	----	----
1918–19	13	1 2 — — — —	$1.50	$3.00
1919–20	14	— — 3 4 5 6	$1.50	$5.00
1920–21	15	1 — — 4 — —	$1.50	$3.00
1921–22	16	— — — — — —	----	----
		Apr. July Oct. Jan.		
1922–23	17	1 — — —	$1.50	$1.50
1923–24	18	— — — —	----	----
1924–25	19	1 2 3 4	$1.50	$5.00
1925–26	20	— — — —	----	----
1926–27	21	— — — —	----	----
1927–28	22	— — — 4	$1.50	$1.50
1928–29	23	— — — 4	$1.50	$1.50
1929–30	24	— 2 3 4	$1.50	$4.50
1930–31	25	1 2 3 4	$1.50	$5.00
1931–32	26	1 2 3 4	$1.50	$5.00
1932–33	27	1 2 3 4	$1.50	$5.00
1933–34	28	1 2 3 4	$1.50	$5.00
1934–35	29	1 2 3 4	$1.50	$5.00
1935–36	30	1 2 3 4	$1.50	$5.00
1936–37	31	1 2 3 4	$1.50	$5.00
		Jan. Apr. July Oct.		
1937	32	1 2 3 & 4	$1.50	$5.00
1938	33	1 2 3 4	$1.50	$5.00
1939	34	1 — 3 4	$1.50	$4.50
1940	35	— — — —	----	----
1941	36	— — — —	----	----
1942	37	1 2 3 4	$1.50	$5.00
1943	38	1 2 3 4	$1.50	$5.00
1944	39	1 2 3 4	$1.50	$5.00
1945	40	1 2 3 4	$1.50	$5.00
1946	41	1 2 3 4	$1.50	$5.00
1947	42	1 2 — 4	$1.50	$4.50
1948	43	1 2 — 4	$1.50	$4.50
1949	44	1 2 3 4	$1.50	$5.00
1950	45	1 2 3 4	$1.75	$6.00
1951	46	1 2 3 4	$1.75	$6.00
1952	47	1 2 3 4	$1.75	$6.00
1953	48	By subscription, $7.00	$2.00	

Only a few issues printed before 1930 are still available.

The price of the Journal of Abnormal and Social Psychology for 1953 has been changed to $7.00 per year and $2.00 per number. From 1950 through 1952 the journal cost $6.00 per year and $1.75 per number. Before 1950, the journal cost $5.00 per year. For incomplete volumes, the price is $1.50 for each available number. For foreign orders, $.25 per volume should be added. The American Psychological Association gives the following discounts:

10% on orders of $50.00 or over for any one journal
20% on orders of $100.00 or over for any one journal
30% on orders of $150.00 or over for any one journal

Orders for available back issues and for current subscriptions should be addressed to

American Psychological Association

1333 Sixteenth Street N. W. Washington 6, D. C.

Published This Month!

A BASIC TEXT FOR COURSES IN GROUP DYNAMICS
A COLLATERAL TEXT FOR SOCIAL PSYCHOLOGY COURSES

GROUP DYNAMICS:
RESEARCH AND THEORY

Edited by Dorwin Cartwright, Professor of Psychology and Director, Research Center for Group Dynamics, University of Michigan; Alvin Zander, Associate Professor of Educational Psychology and Program Director, Research Center for Group Dynamics, University of Michigan.

The facts known about group dynamics as of 1953 . . . important research findings about the way groups function (families, clubs, committees, work-teams, classrooms, military units, etc.); about group leadership; group effectiveness, the setting of group goals; group communications; relations between groups and individuals.

Forty-one chapters by 52 scholars, with six extensive introductions (by the editors) giving over-all views of the subject.

Probably $6.00

Row, Peterson and Company

Evanston, Illinois **White Plains, New York**

[redacted]

The McGRAW-HILL

PERSONALITY and ADJUSTMENT

By WILLIAM E. KAPP, Los Angeles City College, and LEONARD ...
Valley College, Holdrege. Ready in May.

Combining detailed discussions of the scientific approach ... the various areas of living, this text offers a unique ... understanding adjustment and mental hygiene. Psychology, physiology, psychology, mental hygiene, and understanding of the individual.

OBSERVATION. The Basis of Control ...

By E. C. Bentley, The University of Oregon. ... and Anthropology. 142 pages, $2.00

A significant work in applied psychology and a ... of social change. Its new viewpoint and basic general ... generally observational and practical in treatment. ... applied to psychological tests; other backgrounds; ... settings; and the methodology of observing, recording ... series of individuals who are predisposed to accept ...

THE PSYCHOLOGY OF INDUSTRIAL RELATIONS

By C. H. Lawshe, Purdue University. 450 pages, $4.00

Presented in an unusual, psychological and practical manner, ... management students. Those who have had ... and psychology often. Almost completely ... and ... it also is useful to the college student ...

PSYCHOANALYTIC THEORIES OF PER...

By GERALD S. BLUM, University ... Psychology. 320 pages, $3.75

This new text seeks to present ... psychoanalytic writings ... fresh approach to personality ... views in detail and ... the author includes a ... of the ... of existing psychological tests, ... orientation, and evaluation ...

Send for ...

McGRAW-HILL BOOK CO...

330 West 42nd Street